

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 18, 2003, 09:19:25 ; Search time 1703 Seconds

(Without alignments)  
8207.108 Million cell updates/sec

Title: US-10-021-811-35

Perfect score: 863  
Sequence: 1 gcaagagctatcacacac.....aaaaaaaaaaaaaaaaaaaa 863

Scoring table: OLIGO\_NUC  
Gapop 60.0 , Gapext 60.0

Searched: 16154066 seqs, 8097743376 residues

Word size : 60

Total number of hits satisfying chosen parameters: 4

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

EST:\*  
1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estin:\*  
4: em\_estnu:\*  
5: em\_estov:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_hic:\*  
9: gb\_est1:\*  
10: gb\_est2:\*  
11: gb\_hic:\*  
12: gb\_est3:\*  
13: gb\_est4:\*  
14: gb\_est5:\*  
15: em\_estfun:\*  
16: em\_estom:\*  
17: gb\_gss:\*  
18: em\_gss\_hum:\*  
19: em\_gss\_inv:\*  
20: em\_gss\_pln:\*  
21: em\_gss\_vrt:\*  
22: em\_gss\_fun:\*  
23: em\_gss\_mam:\*  
24: em\_gss\_mus:\*  
25: em\_gss\_other:\*  
26: em\_gss\_pro:\*  
27: em\_gss\_rtd:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	562	65.1	562	14	BM732121 sal73b11.
2	370	42.9	586	12	BB820766 GM700012A
3	349	40.4	431	10	AA459279 sh22h08.y
4	285	33.0	450	10	BE058947 sn23b01.y

#### ALIGNMENTS

RESULT 1  
LOCUS BM732121  
DEFINITION sal73b11.y1 Gm-cl061 Glycine max cDNA clone SOYBEAN CLONE ID: Gm-cl061-4678 5' similar to TR:Q39028 Q39028 ATMYB2.1; mRNA sequence.  
ACCESSION BM732121  
VERSION BM732121.1 GI:19053454  
KEYWORDS EST.  
SOURCE soybean.  
ORGANISM Glycine max  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.

REFERENCE 1 (bases 1 to 562)  
Shoemaker,R., Keim,P., Vodkin,L., Erpelting,J., Coryell,V., Khanna,A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers,Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schuk,R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R., Waterston,R. and Wilson,R.  
Public Soybean EST Project  
Unpublished (1999)  
Contact: Shoemaker R/Public Soybean EST Project  
Public Soybean EST Project  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: est@watson.wustl.edu

This clone is available through: Resgen, Invitrogen Corp. 2130 South Memorial Parkway Huntville, AL 35801 For further information call: (800)-533-4363 or contact: ccu@resgen.com web site: www.resgen.com  
Seq primer: -40RP from Glibco  
High quality sequence stop: 421.

#### FEATURES

source

1..562

/organism="Glycine max"

/db\_xref="taxon:3847"

/clone="SOYBEAN CLONE ID: Gm-cl061-4678"

/clone\_lib="Gm-cl061"

/tissue\_type="mature flowers of field grown plants"

/lab\_host="DH10B"

/note="Vector: pluscript II SK+; Site\_1: EcoRI; Site\_2: XhoI; The cDNA library was constructed from mRNA isolated from mature flowers of field grown plants for the cultivar Raiden. Complementary DNA was synthesized from mRNA using a primer consisting of a poly(dT) sequence with a XhoI restriction site. EcoRI adapters were ligated to the blunt-ended cDNA fragments followed by XhoI digestion. The cDNA fragments were directionally cloned into the EcoRI-XhoI restriction site of the pluscript vector. The ligated cDNA fragments were transformed into DH10B host cells (GibcoBRL). This library was constructed in the laboratory of Dr. Randy Shoemaker."

BASE COUNT 187 a 115 c 128 g 132 t  
ORIGIN

Query Match 65.1%; Score 562; DB 14; Length 562;

Best Local Similarity 100.0%; Pred. No. 3.3e-128; Mismatches 0; Indels 0; Gaps 0;

OY	10	CTATCACACACCAAGTCAATGATATAAAACAAGTGTAAAGCTCTCAAGATCTGA	69
DB	1	CTATCACACACCAAGTCAATGATATAAAACAAGTGTAAAGCTCTCAAGATCTGA	60
OY	70	AGTAGAAGAAAGGCGCTTGAGCAATGAGCAAGACCTTGATCTTGATGAACATATTCGAA	129

```

Db 61 AGTGAAGAAAAGGCGCTTGACATGAGAAAGACTTATGATGTAATATTGCAA 120
QY 130 TCATGGGGAAGTGTGGTGAAGCTCTTTGGCCAAAGCTGTGCTCCAAAGCTAACGAAA 189
Db 121 TCATGGGGAAGTGTGGTGAAGCTCTTTGGCCAAAGCTGTGCTCCAAAGCTAACGAAA 180
QY 190 GAGTTGGCGGTAGTGGGTAAATTACCTCCGTCCTGATGTAAGAGGGAATATTAC 249
Db 181 GAGTTGGCGGTAGTGGGTAAATTACCTCCGTCCTGATGTAAGAGGGAATATTAC 240
QY 250 ACCGAGGAACAACTTTGATTTATGAGCTCCAGCCAAAGTGCGGAAACAGTGTCCAA 309
Db 241 ACCGAGGAACAACTTTGATTTATGAGCTCCAGCCAAAGTGCGGAAACAGTGTCCAA 300
QY 310 AATTGCCAAGCATCTACCTGGAAGACTGATTAATGAGATCAAGAACTATTGGAGCAAG 369
Db 301 AATTGCCAAGCATCTACCTGGAAGACTGATTAATGAGATCAAGAACTATTGGAGCAAG 360
QY 370 GATCCAGACACATCAAGCAAGCTGGAACCTTTACAGCAACAGATGTAATATTCTGA 429
Db 361 GATCCAGACACATCAAGCAAGCTGGAACCTTTACAGCAACAGATGTAATATTCTGA 420
QY 430 GATTAATGATCACCAGCTAGCACTAGTGTTCACCATGGCTGAGCCCATGAGAGAT 489
Db 421 GATTAATGATCACCAGCTAGCACTAGTGTTCACCATGGCTGAGCCCATGAGAGAT 480
QY 490 GTATTTCACACCTGTTATCAAGGAATGTTAGACCAATTTTCAAGCTAGTCCCTACAA 549
Db 481 GTATTTCACACCTGTTATCAAGGAATGTTAGACCAATTTTCAAGCTAGTCCCTACAA 540
QY 550 TAATCCGATCAATCCAGTTGT 571
Db 541 TAATCCGATCAATCCAGTTGT 562

```

```

RESULT 2
BE820766/c 586 bp mRNA linear EST 24-MAY-2001
LOCUS GM700012A20H11 Gm-r1070 Glycine max cDNA clone Gm-r1070-4966 3',
DEFINITION mRNA sequence.
ACCESSION BE820766
VERSION BE820766
KEYWORDS EST.
SOURCE soybean.
ORGANISM Glycine max
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
Glycine.
1 (bases 1 to 586)
Vodkin, L., Keim, P., Shoemaker, R., Retzel, E., Khanna, A., Coryell, V.,
Erpelding, J., Raph, C., Shoop, E., Pardins, U., Liu, L. and Lewin, H.
A functional genomics program for soybean (NSF 9872565)
Unpublished (1999)
Other ESTs: AW459279 corresponding to Gm-cl016-5560 (5')
Contact: Vodkin, L.O., PI, A Functional Genomics Program for
Soybean (NSF 9872565)
Lewin, H. A., Director, Keck Center for Comparative and Functional
Genomics
University of Illinois
Edwin R. Madigan Building, 1201 W. Gregory, Urbana, IL 61801, USA
Tel: (217) 244-6147
Fax: (217) 333-4582
Email: l-vodkin@uiuc.edu
This clone is available through: Genome Systems, Inc. 4633 World
Parkway Circle St. Louis, Missouri 63134. For further information
call: (800) 430-0030 or (314) 427-3322 FAX: (888) 919-3324 or (314)
427-3324 or contact: clones@genomesystems.com or info@genome
systems.com web site: www.genomesystems.com
Seq primer: 5'-TTTCTTTTCTTTTCTTTT(A/C/G)-3'.
Location/Qualifiers
1..586
/organism="Glycine max"

```

## FEATURES

Source 1..586  
Location/Qualifiers  
/organism="Glycine max"

```

/db_xref="taxon:3847"
/clone="Gm-r1070-4966"
/clone_id="Gm-r1070"
/note="The library Gm-r1070 is a sequence-driven, rerecked
set of 9,216 clones selected from cDNA libraries from
various tissues and stages of development of soybean that
represent 2,639 sequences from immature cotyledons, 1,770
from immature seed coats, 3,938 from flowers, and 869
from young pods. The 5' ESTs of the source clones from
the different libraries was used to select singletons, or
a representative of each contig, which were rerecked to
form library Gm-r1070. The cDNA clones of the rerecked
Gm-r1070 library were then sequenced at the 3' end. The
contig analysis to select unique genes was performed by
the Laboratory of Ernest Retzel, Center for Computational
Genomics and Bioinformatics, University of Minnesota,
http://www.cbc.umn.edu/ResearchProjects/Soybean/Index.html
. Rerecking was performed by Genome Systems, St. Louis,
http://www.genomesystems.com, and 3' sequencing by the
Keck Center for Comparative and Functional Genomics,
University of Illinois,
http://www.life.uiuc.edu/biotech/keck.html. Note: The
corresponding 5' EST from each clone in the Gm-r1070
library is listed in the 'OTHER EST' field. The detailed
information on the source library for each clone can also
be obtained by referring to the genome systems clone ID of
the original cDNA library that is also listed under
'OTHER EST'."
BASE COUNT 173 a 110 c 100 g 194 t 9 others
ORIGIN

```

```

Query Match 42.9%; Score 370; DB 12; Length 586;
Best Local Similarity 99.6%; Pred. No. 1.9e-81;
Matches 470; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

```

QY 326 CCTGGAAGACTGATTAATGATCAAGAACTATTGGAGGACAGGATCCGAAGACATC 385
Db 583 COTGGAAGACTGATTAATGATCAAGAACTATTGGAGGACAGGATCCGAAGACATC 524
QY 386 AAGCAAGCTGGAAGCTTTACAGCAACAGAGTAGTAATATTCTGAGTAATGATCACC 445
Db 523 AAGCAAGCTGGAAGCTTTACAGCAACAGAGTAGTAATATTCTGAGTAATGATCACC 464
QY 446 GCTAGCACTAGCCATGTTCCACCATGGCTGAGCCCATGAGATGATTTCTCCACCCTGT 505
Db 463 GCTAGCACTAGCCATGTTCCACCATGGCTGAGCCCATGAGATGATTTCTCCACCCTGT 404
QY 506 TATCAAGGAATGTTAGGCCATTTTCACTAGTCCCTCAATTAATCTGATCAATCC 565
Db 403 TATCAAGGAATGTTAGGCCATTTTCACTAGTCCCTCAATTAATCTGATCAATCC 344
QY 566 AGTTGTTGTACCAATGACACACACACATTAACATATTGAGATGAGATAGCTGTCA 625
Db 343 AGTTGTTGTACCAATGACACACACACATTAACATATTGAGATGAGATAGCTGTCA 284
QY 626 ATGCAATTACGAGACGGGTGATTAATATTATCAAGATTAACCTAGTGTGAAGTTCCA 685
Db 283 ATGCAATTACGAGACGGGTGATTAATATTATCAAGATTAACCTAGTGTGAAGTTCCA 224
QY 686 TAAGCTGGAATGCTCTGTGATTAACATATTATGTTGTTTATATATAGTAGTG 745
Db 223 TAAGCTGGAATGCTCTGTGATTAACATATTATGTTGTTTATATATAGTAGTG 164
QY 746 ATGTTGGTTTGGTACCATTAATATTATGTTGTTTATATATAGTAGTG 797
Db 163 ATGTTGGTTTGGTACCATTAATATTATGTTGTTTATATATAGTAGTG 112

```

```

RESULT 3
AW459279 431 bp mRNA linear EST 03-DEC-2001
LOCUS SH22H08.v1 Gm-cl016 Glycine max cDNA clone GENOME SYSTEMS CLONE ID:
DEFINITION Gm-cl016-5560 5' similar to TR:Q92TD5 Q92TD5 PUTATIVE TRANSCRIPTION

```

FACTOR ;, mRNA sequence.  
 AM459279  
 VERSION AM459279.1 GI:7029496  
 KEYWORDS EST.  
 SOURCE soybean.  
 ORGANISM Glycine max  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.  
 1 (bases 1 to 431)  
 REFERENCE  
 1 (bases 1 to 431)  
 AUTHORS Shoemaker,R., Kelm,P., Vodkin,L., Erpelting,J., Corvett,V., Khanna,A., Bolla,B., Matra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers,Y., Person,B., Swaller,T., Gibbons,M., Page,D., Harvey,N., Schurk,R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R., Waterston,R. and Wilson,R.  
 TITLE Public Soybean EST Project  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Shoemaker R/Public Soybean EST Project  
 Public Soybean EST Project  
 Washington University School of Medicine  
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA  
 Tel: 314 286 1800  
 Fax: 314 286 1810  
 Email: est@watson.wustl.edu  
 This clone is available through: Resgen, Invitrogen Corp. 2130 South Memorial Parkway Huntsville, AL 35801 For further information call: (800)-533-4363 or contact via email: cou@resgen.com  
 Insert Length: 723 Std Error: 0.00  
 Seq primer: -40RP from Gibco  
 High quality sequence stop: 411.  
 Location/Qualifiers  
 1..431  
 /organism="Glycine max"  
 /db\_xref="taxon:3847"  
 /clone="GENOME SYSTEMS CLONE ID: Gm-cl016-5560"  
 /clone\_lib="Gm-cl016"  
 /tissue\_type="Immature flowers of field grown plants"  
 /lab\_host="XL10-Gold"  
 /note="Vector: pBluescript II XR; Site:1: EcoRI; Site:2: XhoI; This cDNA library was constructed from mRNA isolated from immature flowers of field grown plants. The cDNA library was prepared using the Stratagene pBluescript II XR library construction kit. Complementary DNA was synthesized from mRNA using a primer consisting of a poly (dT) sequence with a XhoI restriction site. EcoRI adapters were ligated to the blunt-ended cDNA fragments followed by XhoI digestion. The cDNA fragments were directionally cloned into the EcoRI-XhoI restriction site of the pBluescript vector. The ligated cDNA fragments were transformed into XL10-Gold host cells. This library was constructed by Dr. Randy Shoemaker and Dr. John Erpelting."  
 BASE COUNT 145 a 81 c 88 g 117 t  
 ORIGIN  
 Query Match 40.4%; Score 349; DB 10; Length 431;  
 Best Local Similarity 100.0%; Pred. No. 3.3e-76;  
 Matches 349; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 326 CCTGGAAGAGCTGTAATGATCAGACTATTTGGAGCAAGATCCGAAGACATC 385  
 |||||||  
 DB 8 CCTGGAAGAGCTGTAATGATCAGACTATTTGGAGCAAGATCCGAAGACATC 67  
 |||||||  
 QY 386 AAGCAAGCTGAGACTTCGCAACAGATGATTAATCTGAGATTAATGATCCACCA 445  
 |||||||  
 DB 68 AAGCAAGCTGAGACTTCGCAACAGATGATTAATCTGAGATTAATGATCCACCA 127  
 |||||||  
 QY 446 GCTAGACACTAGCCATGTTTCCACCAGCTGCTAGCCCATGAGATGTAATCTCCACCCTGT 505  
 |||||||  
 DB 128 GCTAGACACTAGCCATGTTTCCACCAGCTGCTAGCCCATGAGATGTAATCTCCACCCTGT 187

QY 506 TATCAAGAAATGTTAGACCATTTTCACAGTCAGTCCCTACAAATTTAATCCGATCATCC 565  
 |||||||  
 DB 188 TATCAAGAAATGTTAGACCATTTTCACAGTCAGTCCCTACAAATTTAATCCGATCATCC 247  
 |||||||  
 QY 566 AGTTGTTGACCAATGACAAACAACATTAATGATGAGCAGATGAGCTGTGCA 625  
 |||||||  
 DB 248 AGTTGTTGACCAATGACAAACAACATTAATGATGAGCAGATGAGCTGTGCA 307  
 |||||||  
 QY 626 ATGCATTAACGTAACGGTGATTAATTAATTAATCAAGATTAACCTTAAGTT 674  
 |||||||  
 DB 308 ATGCATTAACGTAACGGTGATTAATTAATTAATCAAGATTAACCTTAAGTT 356  
 |||||||  
 RESULT 4  
 BE058947 450 bp mRNA linear EST 03-DEC-2001  
 LOCUS sn23b01.y1 Gm-cl016 Glycine max cDNA clone GENOME SYSTEMS CLONE ID: Gm-cl016-12290.5, similar to TR:09ZTD5 Q9ZTD5 PUTATIVE  
 DEFINITION TRANSCRIPTION FACTOR ;, mRNA sequence.  
 BE058947  
 BE058947 GI:8403313  
 ACCESSION  
 VERSION BE058947  
 KEYWORDS EST.  
 SOURCE soybean.  
 ORGANISM Glycine max  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.  
 1 (bases 1 to 450)  
 REFERENCE  
 1 (bases 1 to 450)  
 AUTHORS Shoemaker,R., Kelm,P., Vodkin,L., Erpelting,J., Corvett,V., Khanna,A., Bolla,B., Matra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers,Y., Person,B., Swaller,T., Gibbons,M., Page,D., Harvey,N., Schurk,R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R., Waterston,R. and Wilson,R.  
 TITLE Public Soybean EST Project  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Shoemaker R/Public Soybean EST Project  
 Public Soybean EST Project  
 Washington University School of Medicine  
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA  
 Tel: 314 286 1800  
 Fax: 314 286 1810  
 Email: est@watson.wustl.edu  
 This clone is available through: Resgen, Invitrogen Corp. 2130 South Memorial Parkway Huntsville, AL 35801 For further information call: (800)-533-4363 or contact via email: cou@resgen.com  
 Seq primer: -40RP from Gibco  
 High quality sequence stop: 411.  
 Location/Qualifiers  
 1..450  
 /organism="Glycine max"  
 /db\_xref="taxon:3847"  
 /clone="GENOME SYSTEMS CLONE ID: Gm-cl016-12290"  
 /clone\_lib="Gm-cl016"  
 /tissue\_type="Immature flowers of field grown plants"  
 /lab\_host="XL10-Gold"  
 /note="Vector: pBluescript II XR; Site:1: EcoRI; Site:2: XhoI; This cDNA library was constructed from mRNA isolated from immature flowers of field grown plants. The cDNA library was prepared using the Stratagene pBluescript II XR library construction kit. Complementary DNA was synthesized from mRNA using a primer consisting of a poly (dT) sequence with a XhoI restriction site. EcoRI adapters were ligated to the blunt-ended cDNA fragments followed by XhoI digestion. The cDNA fragments were directionally cloned into the EcoRI-XhoI restriction site of the pBluescript vector. The ligated cDNA fragments were transformed into XL10-Gold host cells. This library was constructed by Dr. Randy Shoemaker and Dr. John Erpelting."  
 BASE COUNT 151 a 84 c 90 g 125 t  
 ORIGIN

Query Match 33.0%; Score 285; DB 10; Length 450;  
Best Local Similarity 99.7%; Pred. No. 1.2e-60;  
Matches 335; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 339 ATATGAGATCAGAGACTATTGGAGACAGATCCAGAGACATCAAGCAGCTGAGA 398  
|||||  
Db 36 ATATGAGATCAGAGACTATTGGAGACAGATCCAGAGACATCAAGCAGCTGAGA 95  
|||||

QY 399 ACTTCAGCAGAGAGTAGTATATCTGATTAATGATCAGCAGCTAGCAGCTAGCC 458  
|||||  
Db 96 ACTTCAGCAGAGAGTAGTATATCTGATTAATGATCAGCAGCTAGCAGCTAGCC 155  
|||||

QY 459 ATGTTCCACCATGGCTGAGCCCATGGAGATGATCTCCACCCGTTATCAAGGAATGT 518  
|||||  
Db 156 ATGTTCCACCATGGCTGAGCCCATGGAGATGATCTCCACCCGTTATCAAGGAATGT 215  
|||||

QY 519 TAGAGCCATTTCAACTAGTCCCTACATTAATTCCTGATCAATCCAGTGTGTACCA 578  
|||||  
Db 216 TAGAGCCATTTCAACTAGTCCCTACATTAATTCCTGATCAATCCAGTGTGTACCA 275  
|||||

QY 579 ATGACACACACACATTACTATTGGAGCATGGAGATAGCTGTCATGCAATTACTGA 638  
|||||  
Db 276 ATGACACACACACATTACTATTGGAGCATGGAGATAGCTGTCATGCAATTACTGA 335  
|||||

QY 639 ACGGTGATTAATATTATCAAGATAAAACCTAAGT 674  
|||||  
Db 336 ACGGTGATTAATATTATCAAGATAAAACCTAAGT 371  
|||||

Search completed: February 18, 2003, 10:31:23  
Job time 1705 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 18, 2003, 07:08:59 ; Search time 2178 Seconds

(without alignments)  
11531.553 Million cell updates/sec

Title: US-10-021-811-35

Perfect score: 1 gcacgagcctctacacac.....aaaaaaaaaaaaaaaaaaaa 863

Sequence: 1 gcacgagcctctacacac.....aaaaaaaaaaaaaaaaaaaa 863

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

```

1: gb_da:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_ph:*
7: gb_pl:*
8: gb_pr:*
9: gb_ro:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vl:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rod:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htggo_hum:*
40: em_htggo_mus:*
41: em_htggo_other:*

```

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	382.8	44.4	1022	8	PSMYB26
2	239.4	27.7	910	8	AY086615
3	236.2	27.4	645	8	AF175987
4	215	24.9	1033	8	AB005688
5	208	24.1	1127	8	AY088761
6	194.6	22.5	872	8	AB058642
7	176.6	20.5	1117	8	AF510112
8	175	20.3	1295	8	CP033917
9	163.4	18.9	1423	8	AF262733
10	156.4	18.1	1343	8	AY026332
11	155.6	18.0	1140	8	AF334815
12	154	17.8	817	8	AY133705
13	149.2	17.3	931	8	AY008377
14	149.2	17.3	1024	8	AF062894
15	146.2	16.9	1033	8	AF034133
16	144	16.7	133242	8	AF003140
17	144	16.7	145491	8	AP002883
18	142.2	16.5	771	8	AF272733
19	138	16.0	1753	8	AF411970
20	137.8	16.0	1070	8	AT062743
21	135.2	15.7	1337	8	LETHM16
22	134	15.5	1373	8	LETHM1
23	131	15.2	1341	8	AY099777
24	131	15.2	1364	8	AB005689
25	130.8	15.2	844	8	AT026937
26	130.4	15.1	1373	8	AF214116
27	130	15.1	1293	8	AY062619
28	129.6	15.0	984	8	AF175990
29	128.6	14.9	1038	8	AF249309
30	128.6	14.9	1493	8	AF062916
31	128.4	14.9	1335	8	ATC04GENE
32	128	14.8	1132	8	AY096523
33	128	14.8	1512	8	AY063939
34	127.6	14.8	1262	8	AF371980
35	127.2	14.7	1220	8	AF161711
36	127	14.7	981	8	AF370613
37	127	14.7	1252	8	ATM1X1A
38	126.8	14.7	1331	8	AY087252
39	126.6	14.7	1513	8	AF427146
40	126.6	14.7	1601	8	ZM057002
41	126.6	14.7	1802	8	MZEBPR
42	126.2	14.6	1606	8	AF034132
43	125.2	14.5	318	8	AF474125
44	125.2	14.5	579	8	AF502295
45	124.4	14.4	813	8	AF175993

#### ALIGNMENTS

```

RESULT 1
PSMYB26      1022 bp      mRNA      linear      PLN 09-FEB-1998
LOCUS        PSMYB26
DEFINITION   P.sativum mRNA for Myb-like protein (Myb26).
ACCESSION    Y11105
VERSION      Y11105.1 GI:1841474
KEYWORDS     Myb-like protein; Myb26.
SOURCE       P.sativum
ORGANISM     Pisum sativum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Viciae;
Pisum.
REFERENCE    1 (bases 1 to 1022)
AUTHORS      Uimari,A. and Strommer,J.

```



[illegible]

LOCUS	AF175987	645 bp	mRNA	linear	PLN 30-AUG-2001
DEFINITION	Arabidopsis thaliana putative transcription factor (MYB24) mRNA,				
ACCESSION	complete cds.				
VERSION	AF175987				
KEYWORDS	AF175987.1 GI:5823306				
SOURCE	Arabidopsis thaliana.				
ORGANISM	Arabidopsis thaliana.				
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.				
AUTHORS	1 (bases 1 to 645) Kranz,H.D., Denekamp,M., Greco,R., Jin,H., Leyva,A., Meisner,R.C., Petroni,K., Urzainqui,A., Bevan,M., Martin,C., Smeekens,S., Tonelli,C., Paz-Ares,J., Bevan,M., Martin,C., Smeekens,S., Towards functional characterisation of the members of the R2R3-MYB gene family from Arabidopsis thaliana Plant J 16 (2), 263-276 (1998)				
TITLE					
JOURNAL	Plant J 16 (2), 263-276 (1998)				
MEDLINE	99056848				
PUBMED	9839469				
REFERENCE	2 (bases 1 to 645) Stracke,R., Werber,M., and Weisshaar,B. The R2R3-MYB gene family in Arabidopsis thaliana Curr. Opin. Plant Biol. 4 (5), 447-456 (2001)				
JOURNAL	Curr. Opin. Plant Biol. 4 (5), 447-456 (2001)				
MEDLINE	21481677				
PUBMED	11597504				
REFERENCE	3 (bases 1 to 645) Stracke,R., and Weisshaar,B. Direct Submission				
AUTHORS	Submitted (06-AUG-1999) Dept. Biochemie, Max-Planck-Institut fuer Zuechtungsforshung, Carl-von-Linne-Weg 10, Koeln D-50829, Germany				
JOURNAL					
TITLE					
FEATURES	Location/Qualifiers				
source	1..645 /organism="Arabidopsis thaliana" /db_xref="taxon:3702" /note="ecotype: Columbia"				
gene	1..645				

CDS	BASE COUNT	ORIGIN
/gene="MYB24"		
1..645		
/gene="MYB24"		
/note="R2R3-MYB family member"		
/codon_start=1		
/product="putative transcription factor"		
/protein_id="AAD53092.1"		
/db_xref="GI:5823307"		
/translation="MEKRESSGGSGSGDAEVRKGPWTMEEDLLIINTIANHGECVWNS LAKSAGLKRTRKSCRLRWMLLRPDVNRGNTTPEQLTINLHAKMGNRSKIAKLHP GRDNEIKRFRKIQRYIIKSGETTVRGSSSEPIINHAATTSVMNDQDETMDMYSP TTSVQHASNINQOLATGVNYPESSIMMPLSDQSEQNYNSVYDDLMPMNINYNGN"	219 a 125 c 153 g 148 t	

Query Match	27.4%	Score 236.2	DB 8	Length 645
Best Local Similarity	71.9%	Pred. No.5.8e-41		
Matches 323	Conservative	0	Mismatches 123	Indels 3
			Gaps	1
0Y	56	TCTCAGATCCCTGGAAGTAGAAAAAGGCGCTTGAGCAATAGAAAGAACTGTGATCTTGTATG	115	
Db	34	TCAGAGATATGCAGAGGTGAGAAAAAGGCGCATGAGCATGGAAGAAATTTGATCTCATC	93	
0Y	116	AACATATTGCAATCATATGAGGGAAGGTGTTTGGAACTCTTTGGCCAAAGCTGCTGCTC	175	
Db	94	AATTTATTCGCCAATCATATGTTGAAGGTGTTTGGAACTCTTCGCCAAATCTGCAGAGACTA	153	
0Y	176	AAACCTAACCGGAAAGATGTCGCGCTTAAGGAGGCTTAAATTAACCTCCGCTCGATTTTGA	235	
Db	154	AAAGCACCGGGGAAAAAGTTGCGCGGTTCGGGTGCTGGAACACTACCTCCGACTGATGTGGA	213	
0Y	236	AGAGGGAATATTACACCCGAGGACAACTTTTGTATTATGAGAGCTCCAGCAAGTGGGA	295	
Db	214	CGGGAAATATTCACACCGAAGAACAGCTCAACCATCAATGAACTTCAAGCAAAATGGGGA	273	
0Y	296	AACAGGTGATCCAAATATGCGCAGAGCATCTACCTGGAGAGAGCTGATATGAGATCAAGAC	355	
Db	274	AATAGTGGTCAAAATTTGCAAAAGCATTTTACAGGAAAMACCGACATGAGATTAAGAT	333	
0Y	356	TATTGAGAGACAGAGATCCAGACATCAAGCAAGCTGAGAACTTTCAGCAACAGAGT	415	
Db	334	TTTTGGAGAGACTMAAGATCCAGMAATATCATATCAAGACGGGAAACGACGACCTTGG	393	
0Y	416	AGTATTATTTTCGATTAATATGATACCAAGATGACATGACATGATTTCCACCATGAGCT	475	
Db	394	TCACAAAGCTCCGAGATTTTATTAACCATATGTCGACACAGAGCATGTATGATGATACT	453	
0Y	476	--GAGCCATGAGAGATGATTTCTCCACC	501	
Db	454	CAGAAACCATGATGATATGATTTCTCCAAC	482	
RESULT 4				
AB005888				
LOCUS	AB005888	1033 bp	mRNA	linear
DEFINITION	Arabidopsis thaliana mRNA for ATMYB3, complete cds.			
ACCESSION	AB005888			
VERSION	AB005888.1	GI:2280527		
KEYWORDS	ATMYB3.			
SOURCE	Arabidopsis thaliana (strain:Columbia) cDNA to mRNA.			
ORGANISM	Arabidopsis thaliana			
REFERENCE	1 (sites)			
AUTHORS	Noji,M., Urao,T., Shinozaki,K.Y. and Shinozaki,K.			
TITLE	Molecular cloning of two cDNAs encoding novel myb homologs from Arabidopsis (Accession Nos. AB005888 and AB005889) (PGR8-111)			
JOURNAL	Plant Physiol. 117, 720 (1998)			
REFERENCE	2 (bases 1 to 1033)			
AUTHORS	Noji,M.			
TITLE	Direct Submission			
JOURNAL	Submitted (18-JUL-1997) Masaki Noji, Chiba University, Faculty of			

Pharmaceutical Sciences: Yayoi-cho 1-33, Inage-ku, Chiba, Chiba  
263, Japan (E-mail: mmoji@ep.chiba-u.ac.jp, Tel: +81-43-290-2906,  
Fax: +81-43-290-2905)

FEATURES  
source Location/Qualifiers

1..1033  
/organism="Arabidopsis thaliana"

/strain="Columbia"

gene /db\_xref="taxon:3702"

1..1033

CDS /gene="Atmyb3"

115..795

/gene="Atmyb3"

/codon\_start=1

/product="ATMYB3"

/protein\_id="BA21618.1"

/db\_xref="GI:2280528"

/translation="MEKRGSGSGSSSAEAYRKGPWMEEDLLIINYANHGCV  
WNSLAKSGLKRTGSCRLNLTARDVGRNTPEEQLIMLHKMKRWKIK  
HLRGTDEIKRRCRTRQRIKOSDVTTSVSHHSEINDQASTSHNVCTOD  
QAMETSTPTSTQHTNMEFNTGNTSAAATATVDYFPMVDQTEWTGMDIMS  
SMHLNGN"

BASE COUNT: 362 a 182 c 208 g 281 t

ORIGIN

Query Match

Best Local Similarity 76.7%; Pred. No. 2,1e-36;

Matches 263; Conservative 0; Mismatches 80; Indels 0; Gaps 0;

Score 215; DB 8; Length 1033;

24.9%;

24.9%;

24.9%;

24.9%;

24.9%;

24.9%;

24.9%;

24.9%;

24.9%;

24.9%;

24.9%;

24.9%;

24.9%;

24.9%;

24.9%;

24.9%;

24.9%;

24.9%;

24.9%;

24.9%;

24.9%;

24.9%;

24.9%;

24.9%;

24.9%;

Fieldmann, K.  
Full-Length cDNA from Arabidopsis thaliana  
Unpublished  
3 (bases 1 to 1127)

REFERENCE

Brover, V., Troukhan, M., Alexandrov, N., Lu, Y.-P., Flavell, R. and

Fieldmann, K.

Direct Submission

Submitted (11-MAR-2002) Ceres, Inc. 3007 Malibu Canyon Road,

Malibu, CA 90265, USA

This clone sequence is one of 5,000 Ceres full-length cDNAs made

available to TIGR and Genbank. The following quality assessment of

this set was done by comparison with known proteins: two percent of

the clones are estimated to be 5'-truncated; less than one percent

are 3'-truncated; approximately two percent represent alternative

splice variants, including unsplliced introns and spliced exons; one

percent may contain premature stop codons; five percent may have

frame shifts in a coding region. A sequence is considered to be

5'-truncated if it lacks the translation initiation start (ATG). A

sequence is considered to be 3'-truncated if it lacks the

C-terminal end of the encoded protein. Please note that these cDNA

sequences are derived from the WS or Laer ecotypes and therefore

may contain polymorphisms when compared to sequences from Col-0.

Genet carried out the library production and sequencing of the

full-length clones. Ceres, Inc. carried out the clustering of the

5' sequences, selection of clones, and sequence assembly.

Location/Qualifiers

1..1127

/organism="Arabidopsis thaliana"

/db\_xref="taxon:3702"

/clone="94595"

274..894

/product="putative transcription factor"

/protein\_id="AA67076.1"

/db\_xref="GI:21618026"

/translation="METTKKKRVRKATITSGKEEGYRKGPWMEEDLLIINYANHGCV  
HGBGLNVSVAKSLKRTGSCRLNLTARDVGRNTPEEQLIMLHKMKRWKIK  
SKIAHLRGTDEIKRRCRTRQRIKOSDVTTSVSHHSEINDQASTSHNVCTOD  
AIDTASFQOATTPFNVEQOSNNENYVMDLWVHLNGDHVI"

BASE COUNT: 353 a 203 c 241 g 330 t

ORIGIN

Query Match

Best Local Similarity 77.1%; Pred. No. 6.7e-35;

Matches 253; Conservative 0; Mismatches 75; Indels 0; Gaps 0;

Score 208; DB 8; Length 1127;

24.1%;

24.1%;

24.1%;

24.1%;

Fieldmann, K.  
Full-Length cDNA from Arabidopsis thaliana  
Unpublished  
3 (bases 1 to 1127)

REFERENCE

Brover, V., Troukhan, M., Alexandrov, N., Lu, Y.-P., Flavell, R. and

Fieldmann, K.

Direct Submission

Submitted (11-MAR-2002) Ceres, Inc. 3007 Malibu Canyon Road,

Malibu, CA 90265, USA

This clone sequence is one of 5,000 Ceres full-length cDNAs made

available to TIGR and Genbank. The following quality assessment of

this set was done by comparison with known proteins: two percent of

the clones are estimated to be 5'-truncated; less than one percent

are 3'-truncated; approximately two percent represent alternative

splice variants, including unsplliced introns and spliced exons; one

percent may contain premature stop codons; five percent may have

frame shifts in a coding region. A sequence is considered to be

5'-truncated if it lacks the translation initiation start (ATG). A

sequence is considered to be 3'-truncated if it lacks the

C-terminal end of the encoded protein. Please note that these cDNA

sequences are derived from the WS or Laer ecotypes and therefore

may contain polymorphisms when compared to sequences from Col-0.

Genet carried out the library production and sequencing of the

full-length clones. Ceres, Inc. carried out the clustering of the

5' sequences, selection of clones, and sequence assembly.

Location/Qualifiers

1..1127

/organism="Arabidopsis thaliana"

/db\_xref="taxon:3702"

/clone="94595"

274..894

/product="putative transcription factor"

/protein\_id="AA67076.1"

/db\_xref="GI:21618026"

/translation="METTKKKRVRKATITSGKEEGYRKGPWMEEDLLIINYANHGCV  
HGBGLNVSVAKSLKRTGSCRLNLTARDVGRNTPEEQLIMLHKMKRWKIK  
SKIAHLRGTDEIKRRCRTRQRIKOSDVTTSVSHHSEINDQASTSHNVCTOD  
AIDTASFQOATTPFNVEQOSNNENYVMDLWVHLNGDHVI"

BASE COUNT: 353 a 203 c 241 g 330 t

ORIGIN

Query Match

Best Local Similarity 77.1%; Pred. No. 6.7e-35;

Matches 253; Conservative 0; Mismatches 75; Indels 0; Gaps 0;

Score 208; DB 8; Length 1127;

24.1%;

24.1%;

24.1%;

24.1%;

Fieldmann, K.  
Full-Length cDNA from Arabidopsis thaliana  
Unpublished  
3 (bases 1 to 1127)

REFERENCE

Brover, V., Troukhan, M., Alexandrov, N., Lu, Y.-P., Flavell, R. and

Fieldmann, K.

Direct Submission

Submitted (11-MAR-2002) Ceres, Inc. 3007 Malibu Canyon Road,

Malibu, CA 90265, USA

This clone sequence is one of 5,000 Ceres full-length cDNAs made

available to TIGR and Genbank. The following quality assessment of

this set was done by comparison with known proteins: two percent of

the clones are estimated to be 5'-truncated; less than one percent

are 3'-truncated; approximately two percent represent alternative

splice variants, including unsplliced introns and spliced exons; one

percent may contain premature stop codons; five percent may have

frame shifts in a coding region. A sequence is considered to be

5'-truncated if it lacks the translation initiation start (ATG). A

sequence is considered to be 3'-truncated if it lacks the

C-terminal end of the encoded protein. Please note that these cDNA

sequences are derived from the WS or Laer ecotypes and therefore

may contain polymorphisms when compared to sequences from Col-0.

Genet carried out the library production and sequencing of the

full-length clones. Ceres, Inc. carried out the clustering of the

5' sequences, selection of clones, and sequence assembly.

Location/Qualifiers

1..1127

/organism="Arabidopsis thaliana"

/db\_xref="taxon:3702"

/clone="94595"

274..894

/product="putative transcription factor"

/protein\_id="AA67076.1"

/db\_xref="GI:21618026"

/translation="METTKKKRVRKATITSGKEEGYRKGPWMEEDLLIINYANHGCV  
HGBGLNVSVAKSLKRTGSCRLNLTARDVGRNTPEEQLIMLHKMKRWKIK  
SKIAHLRGTDEIKRRCRTRQRIKOSDVTTSVSHHSEINDQASTSHNVCTOD  
AIDTASFQOATTPFNVEQOSNNENYVMDLWVHLNGDHVI"

BASE COUNT: 353 a 203 c 241 g 330 t

ORIGIN

Query Match

Best Local Similarity 77.1%; Pred. No. 6.7e-35;

Matches 253; Conservative 0; Mismatches 75; Indels 0; Gaps 0;

Score 208; DB 8; Length 1127;

24.1%;

24.1%;

24.1%;

24.1%;

Fieldmann, K.  
Full-Length cDNA from Arabidopsis thaliana  
Unpublished  
3 (bases 1 to 1127)

REFERENCE

Brover, V., Troukhan, M., Alexandrov, N., Lu, Y.-P., Flavell, R. and

Fieldmann, K.

Direct Submission

Submitted (11-MAR-2002) Ceres, Inc. 3007 Malibu Canyon Road,

Malibu, CA 90265, USA

This clone sequence is one of 5,000 Ceres full-length cDNAs made

available to TIGR and Genbank. The following quality assessment of

this set was done by comparison with known proteins: two percent of

the clones are estimated to be 5'-truncated; less than one percent

are 3'-truncated; approximately two percent represent alternative

splice variants, including unsplliced introns and spliced exons; one

percent may contain premature stop codons; five percent may have

frame shifts in a coding region. A sequence is considered to be

5'-truncated if it lacks the translation initiation start (ATG). A

sequence is considered to be 3'-truncated if it lacks the

C-terminal end of the encoded protein. Please note that these cDNA

sequences are derived from the WS or Laer ecotypes and therefore

may contain polymorphisms when compared to sequences from Col-0.

Genet carried out the library production and sequencing of the

full-length clones. Ceres, Inc. carried out the clustering of the

5' sequences, selection of clones, and sequence assembly.

Location/Qualifiers

1..1127

/organism="Arabidopsis thaliana"

/db\_xref="taxon:3702"

/clone="94595"

274..894

/product="putative transcription factor"

/protein\_id="AA67076.1"

/db\_xref="GI:21618026"

/translation="METTKKKRVRKATITSGKEEGYRKGPWMEEDLLIINYANHGCV  
HGBGLNVSVAKSLKRTGSCRLNLTARDVGRNTPEEQLIMLHKMKRWKIK  
SKIAHLRGTDEIKRRCRTRQRIKOSDVTTSVSHHSEINDQASTSHNVCTOD  
AIDTASFQOATTPFNVEQOSNNENYVMDLWVHLNGDHVI"

BASE COUNT: 353 a 203 c 241 g 330 t

ORIGIN

Query Match

Best Local Similarity 77.1%; Pred. No. 6.7e-35;

Matches 253; Conservative 0; Mismatches 75; Indels 0; Gaps 0;

Score 208; DB 8; Length 1127;

24.1%;

24.1%;

24.1%;

24.1%;

Fieldmann, K.  
Full-Length cDNA from Arabidopsis thaliana  
Unpublished  
3 (bases 1 to 1127)

REFERENCE

Brover, V., Troukhan, M., Alexandrov, N., Lu, Y.-P., Flavell, R. and

Fieldmann, K.

Direct Submission

Submitted (11-MAR-2002) Ceres, Inc. 3007 Malibu Canyon Road,

Malibu, CA 90265, USA

This clone sequence is one of 5,000 Ceres full-length cDNAs made

available to TIGR and Genbank. The following quality assessment of

this set was done by comparison with known proteins: two percent of

the clones are estimated to be 5'-truncated; less than one percent

are 3'-truncated; approximately two percent represent alternative

splice variants, including unsplliced introns and spliced exons; one

percent may contain premature stop codons; five percent may have

frame shifts in a coding region. A sequence is considered to be

5'-truncated if it lacks the translation initiation start (ATG). A

sequence is considered to be 3'-truncated if it lacks the

C-terminal end of the encoded protein. Please note that these cDNA

sequences are derived from the WS or Laer ecotypes and therefore

may contain polymorphisms when compared to sequences from Col-0.

Genet carried out the library production and sequencing of the

full-length clones. Ceres, Inc. carried out the clustering of the

5' sequences, selection of clones, and sequence assembly.

Location/Qualifiers

1..1127



RESULT 7	AF510112	1117 bp	mRNA	linear	PLN 26-MAY-2002
LOCUS	AF510112				
DEFINITION	Craterostigma plantagineum MYB transcription factor (MYB10) mRNA,				
ACCESSION	AF510112				
VERSION	AF510112.1				
KEYWORDS					
SOURCE	Craterostigma plantagineum.				
ORGANISM	Craterostigma plantagineum				
REFERENCE	1 (bases 1 to 1117)				
AUTHORS	Villalobos,M.A., Bartels,D. and Iturriaga,G.				
TITLE	The Craterostigma plantagineum CPMYB10 gene enhances stress tolerance to Arabidopsis transgenic plants				
JOURNAL	Unpublished				
REFERENCE	2 (bases 1 to 1117)				
AUTHORS	Villalobos,M.A. and Iturriaga,G.				
TITLE	Direct Submission				
JOURNAL	Submitted (07-MAY-2002) Plant Molecular Biology, Institute of Biotechnology, Universidad Ave. 2001, Cuernavaca, Mor 62210, Mexico				
FEATURES	location/Qualifiers				
source	1..1117				
gene	/organism="Craterostigma plantagineum"				
	/db_xref="taxon:4153"				
	1..1117				
	/gene="MYB10"				
	/note="CPMYB10"				
	79..1080				
	/gene="MYB10"				
	/codon_start=1				
	/product="MYB transcription factor"				
	/protein_id="AA043912.1"				
	/db_xref="GI:21215202"				
	/translation="MNQOQVSKNNQVNNCCEDDDSSDLRRPMTVDEPFLINTI AHHEGRMNSLARRPAGCIKRTKSGSRLMLNLTLRDVRNGTITTEQLLILHSRWGN RWSIAOHLPGRTDNEIKNYWRTKRVQVHQAOLKCDVNSKFKDMRYLMPRLVERIO ASATTDGAPPAVAVSPSSAMNTACYSANMAAGDHRQFMIPQYVATTTTHNSMI AOEWSIVASSESGSLSELTAEANTANHRVINGADHOOIDSTSYDMONCAVGN NGNSDOLGMGMDADRRSNEQMMMTDIDVVDNGSSDDDNMLNVDVWFLQEPSSCF"				
BASE COUNT	329 a 265 c 271 g 252 t				
ORIGIN					
Query Match	20.5%; Score 176.6; DB 8; Length 1117;				
Best Local Similarity	71.3%; Pred. No. 4.2e-28;				
Matches 233; Conservative 0; Mismatches 94; Indels 0; Gaps 0;					
64	TCCTGTAAGTGAAGAAAGGCGCTTGACACATGAGAGACTTGATCTTGATGAATATAT 123				
150	TTTCGACCTCGAGAGAGGACCTTGACCTGTGATGAGAACTTCACTACATCACTACAT 209				
124	TGCAAAATCATGGGGAGGTGTTGGAACCTCTTTGGCCCAACAGCTGTGTGCTCAAAAGTAA 183				
210	CGCTCACCATGCGCGGAGGAGAGACATGGAACCTCTCTTGCACGTTTCCGTGGCTGAAACGAAAC 269				
184	CGGAAAGAGTGGCGGCTTAAGGTGGCTAAATTAACCTCCGCTCGTGTGTAGAAAGAGGAA 243				
270	TGGAAGAGAGTGCAGATTTGAGATGGTTGAACACTACTTGAAGACCCAGACCTTCTCGCAGAGAA 329				
244	TATTATACCCGAGAGAACACTTTTGATTTATGAGAGCTCCAGCCAAAGTGGGGAAGAGAGT 303				
330	TATCACTCGAAGAACGCTTCTCATCTCTGAACCTCATTCACGTTTGAGGCGCAATAGGTG 389				
304	GGCAAAATTTGCCAAGCTTACCGCGGAGAGAGCTGATTAATGAGATCAAGACTAATTTGAG 363				
390	GTCGAAGATTGCCCAACACTTTTGGCTGGAGAGAGGAGACATGATGAATAAAGAACTACTGAG 449				
364	GACAAAGATCCGAGAGACATCAAGCA 390				

Db 450 AACAGGGTCCAAAACATGCCAACA 476

## RESULT 8

CPU33917 1295 bp mRNA linear PLN 17-Apr-2001

LOCUS Craterostigma plantagineum myb-related transcription factor (cpm)

DEFINITION mRNA, complete cds.

ACCESSION U33917

VERSION U33917.1 GI:1002799

KEYWORDS Craterostigma plantagineum.

SOURCE Craterostigma plantagineum

ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Asteridae; eusterids I; Lamiales; Lamiales incertae sedis; Torenidae; Craterostigma.

REFERENCE 1 (bases 1 to 1295)

AUTHORS Iturriza, G., Leyns, L., Villegas, A., Charalbeh, R., Salami, F. and Bartels, D.

TITLE A family of novel myb-related genes from the resurrection plant

Craterostigma plantagineum are specifically expressed in callus and roots in response to ABA or desiccation

Plant Mol. Biol. 32 (4), 707-716 (1996)

JOURNAL MEDLINE 97134962

REFERENCE 2 (bases 1 to 1295)

AUTHORS Iturriza, G.

TITLE Direct Submission

Submitted (14-AUG-1995) Gabriel Iturriza, Plant Molecular Biology, Instituto de Biotecnologia, Av. Universidad #2001, Cuernavaca, MOR

JOURNAL 62210, Mexico

FEATURES Location/Qualifiers

1..1295

/organism="Craterostigma plantagineum"

/db\_xref="taxon:4153"

1..1295

gene /gene="cpm7"

125..1132

/gene="cpm7"

/note="putative DNA-binding protein; Description: myb-like

gene; myb-related transcription factor"

/product="Cpm7"

/codon\_start=1

/protein\_id="AAB58314.1"

/db\_xref="GI:1002800"

/translation="MNOQYVYKSKKKKINSEDDSSDLRGPTVDEFTLINTIA

HHGGRRNSIARFAGLKRKSCRLMINTLRDVRGNTLLEOILLILFHSRMGR

WSKTAOHLPTDNEIKNYRTYRQYKHAOLKCDVNSKQPKDRIYLMRPLVERIOR

AAITDDGAPPLASASPSSAMNTACTSTMAAGDRHROPMLPOTYAMTTHSNLITA

QENSTYVASESEFGLSSSELTEANYANHOVINGADHQIDSSSTTSYGMQNDPAINT

VGNNGNSDIQMGMSVDDRSNEDMMMDVDNCGSSDKDNLIMVNDVWFLDOOFSSC

F"

BASE COUNT 405 a 286 c 279 g 325 t

ORIGIN

Query Match 20.3%; Score 175; DB 8; Length 1295;

Best Local Similarity 70.9%; Pred. No. 9.1e-26;

Matches 232; Conservative 0; Mismatches 95; Indels 0; Gaps 0;

QY 64 TCCTGAGTGAAGAAAGGCTTGACAAATGGAAGAACTGATGTAATAT 123

DB 193 TTCGACCTCGAAGAGAGCGGTGACTGTGATGAAGACTCACTCACTACAT 252

QY 124 TGCAAAATATGGGAAGAGTGTGGAACCTCTTGCCAAAGCTCGTCTCAAGCTAA 183

DB 253 CGCTACCATGCGAAGAGAGAGTGTGGAACCTCTTGCCAAAGCTCGTCTCAAGCTAA 312

QY 184 TGAAGAAGCTTCCGCGCTAAGCTGCTAAATTAACCTCGCTGCTGTAAGAGAGGAA 243

DB 313 TGAAGAAGCTGCAAGATGATGTTAACTACTTGAAGCCGATGTGCTGAGGAAA 372

QY 244 TATTACACCCGAGAACAACTTTGATTATGAGCTCCACGAAAGTGGGAAACAGGTG 303

DB 373 TATCACACTCGAAGAACACACTCTCTCATTCCTGATTCCTGCGGCAATAGCTG 432

QY 304 GTCCAAATTTGCCAGACACTTACCTGGAAGACTGATATGATCAATCAATATTTAGG 363

DB 433 GTGGAAGATTGCCACACATTGCTGGAAGAGCAGACAAATGATTAAGAACTACTGGAG 492

QY 364 GACACGATCCAGACGACATCAAGCA 390

DB 493 AACGAGGTCCAAAACATGCCAACA 519

RESULT 9

AF262733

LOCUS Arabidopsis thaliana putative transcription factor MYB108 (MYB108)

DEFINITION mRNA, complete cds.

ACCESSION AF262733

VERSION AF262733.2 GI:15375290

KEYWORDS Arabidopsis thaliana.

SOURCE Arabidopsis thaliana

ORGANISM Arabidopsis thaliana

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

REFERENCE 1 (bases 1 to 1423)

AUTHORS Stracke, R., Werber, M. and Weisshaar, B.

TITLE The R2R3-MYB gene family in Arabidopsis thaliana

Cur. Opin. Plant Biol. 4 (5), 447-456 (2001)

JOURNAL MEDLINE 21481677

REFERENCE 2 (bases 1 to 1423)

AUTHORS Stracke, R. and Weisshaar, B.

TITLE Direct Submission

Submitted (02-MAY-2000) Blochemie, Max-Planck-Institut fuer

Zuechtungsforchung, Carl-von-Linne-Weg 10, Koeln D-50829, Germany

3 (bases 1 to 1423)

REFERENCE 3 (bases 1 to 1423)

AUTHORS Stracke, R. and Weisshaar, B.

TITLE Direct Submission

Submitted (29-AUG-2001) Dept. Plant Breeding and Yield Physiology,

Max-Planck-Institut fuer Zuechtungsforchung, Carl-von-Linne-Weg

10, Koeln D-50829, Germany

Sequence update by submitter

On Aug 30, 2001 this sequence version replaced gi:8101955.

FEATURES Location/Qualifiers

1..1423

/organism="Arabidopsis thaliana"

/db\_xref="taxon:3702"

/chromosome="3"

/note="ecotype: Columbia"

1..1423

/gene="MYB108"

1..154

/gene="MYB108"

5

/gene="MYB108"

/note="alternative transcription start site"

155..1126

/gene="MYB108"

/note="At3g06490; member of the R2R3-MYB family"

/product="putative transcription factor MYB108"

/codon\_start=1

/protein\_id="AAF72668.1"

/db\_xref="GI:8101956"

/translation="MDEGRSLKNNMEDMDLKRGPWTAEDFRKLNMYIATNGEGR

NLSRCAGLQRTGKSCRLRWLNYLRDVRGNITLLEQLILFHSRWGNWSKIAOY

LEGRDNEIKNYRTYRQYKHAOLKCDVNSKQPKDRIYLMRPLVERIORWSASAA

AAATTTTGTGAGTSSCITTSNNOFNMYDYNNNMGOFCVMGNNNDYTPENSVAV

SPASDLTETYSAPNDPEYSGOMKNSYTPPOINTVSSDLPEDFNSGLDEDTLTM

QEOSLWSFENINGAASSSDSLMNIIGETDEDFWFLQDQDQGNNGSF"

1127..1423

/gene="MYB108"

3' UTR

BASE COUNT 498 a 273 c 282 g 370 t  
ORIGIN  
Query Match 18.9%; Score 163.4; DB 8; Length 1423;  
Best Local Similarity 67.4%; Pred. No. 2.9e-25;  
Matches 230; Conservative 0; Mismatches 111; Indels 0; Gaps 0;  
QY 50 MAGAGCTCCAGAGCTCTGAGTGAAGAAAAGGCGCTTGACATGAGAGAGAGCTTGATC 109  
DB 188 AACATGGAAGACGAGATGACCTAAAGAGAGAGCGCTGAGACCTGCTGAAGAGATTAAAG 247  
QY 110 TTATGATACATATTTGCAATCATCTGGGGAAGGTTGGAACTCTTTGGCCAAAGCTGCT 169  
DB 248 CTCATGAATTAATCAATCTCTAATGAGAGAGCGCTGGAACCTCTTCTCGTTGCGGCC 307  
QY 170 GGTCTCAAAAGCTAAGGAAAGAGCTCCGCTAAGGTGGCTAAATTAACCTCCGCTGAT 229  
DB 308 GGGCTCCAAAGCAGCGGTAAGAGCTTAAGCTAAGTGGTTAACTATCTCCGCCCTGAC 367  
QY 230 GTTAGAAGAGGAAATATACACCGAGAGAACACTTTGATTATGAGAGCTCCAGCGAAAG 289  
DB 368 GTCCGCGGTGGAACAATCACTTGAAGAACAACTCTGATVCCCTGCACTTCACTTCCCGT 427  
QY 290 TGGGGAACAGGTGTGTCCAAAATTCGCCAGCATCTACCTGGAAGAGCTGATATGATC 349  
DB 428 TGGGGAATATGATGTGTCAAATAATGCAATATTTACCGGGAAGACGACAAAGATC 487  
QY 350 AAGCACTATTGAGAGACAAAGATCCAGAAAGACATCAAGCA 390  
DB 488 AAGAACTACTGAGAGAGCGGGGTGCAAAAGCATGCGAAACA 528  
RESULT 10  
AY026332 1343 bp mRNA linear PLN 30-JAN-2002  
LOCUS Oryza sativa Myb transcription factor JAMYb mRNA, complete cds.  
DEFINITION AY026332 GI:13177577  
VERSION AY026332.1  
KEYWORDS  
SOURCE Oryza sativa (japonica cultivar-group).  
ORGANISM Oryza sativa (japonica cultivar-group).  
REFERENCE 1 (bases 1 to 1343)  
AUTHORS Lee, M.-W., Qi, M., and Yang, Y.  
TITLE A novel jasmonic acid-inducible rice myb gene associates with fungal infection and host cell death  
JOURNAL Mol. Plant Microbe Interact. 14 (4), 527-535 (2001)  
MEDLINE 21204658  
PUBMED 11310740  
REFERENCE 2 (bases 1 to 1343)  
AUTHORS Lee, M.-W. and Yang, Y.  
TITLE Direct Submission  
JOURNAL Submitted (29-JAN-2001) Plant Pathology, University of Arkansas, 217 Plant Science Bldg., Fayetteville, AR 72701, USA  
FEATURES  
source  
1. 1343  
/organism="Oryza sativa (japonica cultivar-group)"  
/cultivar="Drew"  
/db\_xref="taxon:39947"  
127..984  
/note="related to host cell death and defense responses; induced by jasmonic acid, wounding, or infection of rice blast fungus, but not by salicylic acid or abscisic acid"  
/codon\_start=1  
/product="Myb transcription factor JAMYb"  
/protein\_id="AAK09883.1"  
/translation="MEMVLPRTSHHPVPEQEOEAELSAELRGPMTVDEDLTLIN YISDHGEGRMALARAAGLKRKSGRLKMLNLPDYKRCGNFTAEPLLILIDHSRM GNRSKIAOHLPGRFDNEIKNRRFRVCKHAKLNCQVNSKRFKAMATYLMKPRLAER IHARAAYDDSDSDYNNDLSCVSGVTMAIVANCFGSGSPKMTSSSSDSFSSDLK"

BASE COUNT 367 a 298 c 337 g 341 t  
ORIGIN  
Query Match 18.1%; Score 156.4; DB 8; Length 1343;  
Best Local Similarity 67.5%; Pred. No. 9.6e-24;  
Matches 220; Conservative 0; Mismatches 106; Indels 0; Gaps 0;  
QY 66 CTGAAGTGAAGAAAGGCGCTTGACATGAGAAAGAGCTTATCTGTATGACATATATG 125  
DB 206 CTGAGCTCCGGAGAGGCGCTTGACCTGACAGAGAGAGCTCACTCAATTAATCACT 265  
QY 126 CAATTCATGGGGAAGGTGTTGGAACTCTTGCCAAAGCTGCTGCTCAAGTAAG 185  
DB 266 CTGATCAGCGGAGGCGCGCTGGAAGAGACACACGCGCGCGCTGGAAGAGAGCTG 325  
QY 186 GAAAGAGTTGCCGGCTAAGGTGGCTAAATTAATTAATCTGCTGATGTTAAGAGGGAATA 245  
DB 326 GGAAGAGCTGCGCGCTGCGTGAACATATCTCCGCGGATGTGAAGCGCGCAACT 385  
QY 246 TTACACCGGAGAACACTTTGATTATGAGAGCTCCAGCAAGTGGGGAAGAGCTGCT 305  
DB 386 TCACCGAGAGAGAGAGCTGCTCATCTGACCTCACTCCGATGGGCAACCGATGT 445  
QY 306 CCMAATTTGCCAAGCACTTACCTGGAAGAGCTGATATGATGATCAAGAACTATTGAGAGA 365  
DB 446 CCAACATGAGCAACAATTTTCCTGCGAGAGACCGAACAAGATCAAGAACTACTGAGAGA 505  
QY 366 CAAGATCCAGAGACATCAAGCAA 391  
DB 506 CCAGAGTGCMAAAGCATGCCAAGCAA 531  
RESULT 11  
AF334815 1140 bp mRNA linear PLN 30-AUG-2001  
LOCUS Arabidopsis thaliana putative transcription factor (MYB116) mRNA,  
DEFINITION AF334815  
VERSION AF334815  
KEYWORDS  
SOURCE Arabidopsis thaliana.  
ORGANISM Arabidopsis thaliana.  
REFERENCE 1 (bases 1 to 1140)  
AUTHORS Stracke, R., Werder, M., and Weisshaar, B.  
TITLE The R2R3-MYB gene family in Arabidopsis thaliana  
JOURNAL Curr. Opin. Plant Biol. 4 (5), 447-456 (2001)  
MEDLINE 21481677  
PUBMED 11597504  
REFERENCE 2 (bases 1 to 1140)  
AUTHORS Stracke, R. and Weisshaar, B.  
TITLE Direct Submission  
JOURNAL Submitted (10-JAN-2001) Dept. Plant Breeding and Yield Physiology, Max-Planck-Institut fuer Zuechtungsforschung, Carl-von-Linne-Weg 10, Koeln D-50829, Germany  
REFERENCE 3 (bases 1 to 1140)  
AUTHORS Stracke, R. and Weisshaar, B.  
TITLE Direct Submission  
JOURNAL Submitted (29-AUG-2001) Dept. Plant Breeding and Yield Physiology, Max-Planck-Institut fuer Zuechtungsforschung, Carl-von-Linne-Weg 10, Koeln D-50829, Germany  
REMARK  
COMMENT Sequence update by submitter  
FEATURES  
source  
1. 1140  
/organism="Arabidopsis thaliana"  
/db\_xref="taxon:3702"  
/note="ecotype: Columbia"  
1..1140  
gene

5'UTR /gene="MYB116"  
1..129  
/gene="MYB116"  
130..981  
CDS /gene="MYB116"  
/note="At1g25340; R2R3-MYB family member"  
/codon\_start=1  
/product="putative transcription factor MYB116"  
/protein\_id="AAK25748.2"  
/db\_xref="GI:15375295"

3'UTR  
982..1140  
/gene="MYB116"

BASE COUNT 397 a 202 c 222 g 319 t  
ORIGIN

Query Match 18.0%; Score 155.6; DB 8; Length 1140;  
Best Local Similarity 67.7%; Pred. No. 1.5e-23;  
Matches 218; Conservative 0; Mismatches 104; Indels 0; Gaps 0;

60 AAGATCTGAAAGTGAAGAAAGGCGCTTGACATGGAAGAAAGCTTGATCTTGATGAAC 119  
170 AGGATGACAGAGAGAGAAAGGCGCTTGACATGGAAGAAAGCTTGATCTTGATGAAC 229  
120 AATATGCAATATCATGGGAGAGTGTGGAACTCTTGGCCAAAGCTGCTGCTCAAC 179  
230 AATATGCAATATCATGGGAGAGTGTGGAACTCTTGGCCAAAGCTGCTGCTCAAC 289  
180 GTAACGAAAGAGTGTGGAGTGTGGTAAATTAATTAATTAATTAATTAATTAATTA 239  
290 GAGCAGAGAAAGTGTGGAGTGTGGTAAATTAATTAATTAATTAATTAATTAATTA 349  
240 GGAATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 299  
350 GGAATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 409  
300 GGTGCTCAAAATTCGCAAGATCTGATGAGTGAAGTGAATGAATGAATGAATGAAT 359  
410 GGTGCTCAAAATTCGCAAGATCTGATGAGTGAAGTGAATGAATGAATGAATGAAT 469  
360 GGAGAGCAAGATTCGCAAGATTCGCAAGATTCGCAAGATTCGCAAGATTCGCAAG 381  
470 GGAGAGCAAGATTCGCAAGATTCGCAAGATTCGCAAGATTCGCAAGATTCGCAAG 491

RESULT 12  
AT133705 817 bp mRNA linear PLN 07-AUG-2002  
LOCUS Arabidopsis thaliana clone C103160 putative myb protein (At4g13480)  
DEFINITION mRNA, complete cds.  
ACCESSION AY133705.1 GI:22136639  
VERSION AY133705.1  
KEYWORDS FTI-CDNA.  
SOURCE Arabidopsis thaliana.  
ORGANISM Arabidopsis thaliana.  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;  
Rosidae; eustosids II; Brassicales; Brassicaceae; Arabidopsis.

REFERENCE  
AUTHORS Yamada,K., Chan,M.M., Chang,C.H., Dale,J.M., Deng,J.M., Hsuan,V.W.,  
Lee,J.M., Quach,H.L., Tang,C.C., Toriumi,M., Wu,H.C., Yu,G.,  
Bower,L., Chen,H., Cheuk,R., Jones,T., Karlin-Neumann,G., Kim,C.,  
Lam,B., Lin,J., Miranda,M., Nguyen,M., Palm,C.J., Shinn,P.,  
Southwick,A., Davis,R.W., Ecker,J.R. and Theologis,A.  
Arabidopsis Open Reading Frame (ORF) Clones

TITLE  
JOURNAL Arabidopsis Open Reading Frame (ORF) Clones  
REFERENCE 2 (bases 1 to 817)  
AUTHORS Yamada,K., Chan,M.M., Chang,C.H., Dale,J.M., Deng,J.M., Hsuan,V.W.,

Lee,J.M., Quach,H.L., Tang,C.C., Toriumi,M., Wu,H.C., Yu,G.,  
Bower,L., Chen,H., Cheuk,R., Jones,T., Karlin-Neumann,G., Kim,C.,  
Lam,B., Lin,J., Miranda,M., Nguyen,M., Palm,C.J., Shinn,P.,  
Southwick,A., Davis,R.W., Ecker,J.R. and Theologis,A.  
Direct Submission  
Submitted (17-JUL-2002) Plant Gene Expression Center, 800 Buchanan  
Street, Albany, CA 94710, USA  
location/Qualifiers  
1..817  
/organism="Arabidopsis thaliana"  
/db\_xref="taxon:3702"  
/chromosome="4"  
/clone="C103160"  
/note="This clone is in PUNI 51.  
ecotype: Columbia"  
1..817  
/gene="At4g13480"  
1..786  
/gene="At4g13480"  
/codon\_start=1  
/evidence="experimental"  
/product="putative myb protein"  
/protein\_id="AA091639.1"  
/db\_xref="GI:22136640"

gene  
1..817  
/gene="At4g13480"  
1..786  
/gene="At4g13480"  
CDS

BASE COUNT 264 a 149 c 199 g 205 t  
ORIGIN

Query Match 17.8%; Score 154; DB 8; Length 817;  
Best Local Similarity 63.5%; Pred. No. 3.4e-23;  
Matches 235; Conservative 0; Mismatches 135; Indels 0; Gaps 0;

73 GAGAAAGGCGCTTGACATGGAAGAGAGTGAATGATGATGATGATGATGATGATGAT 132  
18 GAGAAAGGCGCTTGACATGGAAGAGAGTGAATGATGATGATGATGATGATGATGAT 77  
133 TGGGGAAGGCTGTGGAGTGTGGAGTGTGGAGTGTGGAGTGTGGAGTGTGGAGTGTGG 192  
78 CGGTGAAGGCTGTGGAGTGTGGAGTGTGGAGTGTGGAGTGTGGAGTGTGGAGTGTGG 137  
193 TTGCGGCTAAGTGTGCTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 252  
138 CTGACAGCTAAGTGTGCTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 197  
253 CGAGGAAGAGTGTGCTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 312  
198 AATGAAGAGAGTGTGCTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 257  
313 TGGCAAGCATCTGGAAGAGAGTGTGCTAAATTAATTAATTAATTAATTAATTAATTAAT 372  
258 TGCACGATGTTTACAGGAAGAGAGTGTGCTAAATTAATTAATTAATTAATTAATTAATTA 317  
373 CCAGGAAGCATCTGGAAGAGAGTGTGCTAAATTAATTAATTAATTAATTAATTAATTAAT 432  
318 CAAGGAAGAGAGTGTGCTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 377  
433 AATGATCAC 442  
Db 378 AAGCAACAC 387

RESULT 13  
LOCUS AY008377 931 bp mRNA linear PLN 30-AUG-2001  
DEFINITION Arabidopsis thaliana putative transcription factor MYB112 mRNA,  
complete cds.  
ACCESSION AY008377

```

VERSION      AY008377.2  GI:15375307
KEYWORDS
SOURCE       Arabidopsis thaliana.
ORGANISM     Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
              Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
              Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsi.
REFERENCE    1 (bases 1 to 931)
AUTHORS      Stracke,R., Werber,M. and Weisshaar,B.
JOURNAL      The R2R3-MYB gene family in Arabidopsis thaliana
MEDLINE      21481677
PUBMED       11597504
REFERENCE    2 (bases 1 to 931)
AUTHORS      Stracke,R. and Weisshaar,B.
JOURNAL      Direct Submission
TITLE        Submitted (29-SEP-2000) Dept. Biochemie, Max-Planck-Institut fuer
              Zuechtungsforchung, Carl-von-Linne-Weg 10, Koeln D-50829, Germany
REFERENCE    3 (bases 1 to 931)
AUTHORS      Stracke,R. and Weisshaar,B.
JOURNAL      Submitted (29-AUG-2001) Dept. Plant Breeding and Yield Physiology,
              Max-Planck-Institut fuer Zuechtungsforchung, Carl-von-Linne-Weg
              10, Koeln D-50829, Germany
REMARK       Sequence update by submitter
COMMENT      On Aug 30, 2001 this sequence version replaced gi:11641119.
FEATURES
  source      1..931
              location/Qualifiers
              /organism="Arabidopsis thaliana"
              /db_xref="taxon:3702"
              /chromosome="1"
              /note="ecotype: Columbia"
              1..83
              6
              /note="alternative transcription start site"
              84..815
              /note="Atlg48000: R2R3-MYB family member"
              /codon_start=1
              /product="putative transcription factor MYB112"
              /protein_id="AAG38379.2"
              /db_xref="GI:15375308"
              /translation="MNISRTFANCKTLINHKKEVEVEKKEIEIRGPMVEEDMK
              LVSYSILHGEKWNLSRSAGLNRTGKSRMLANLRPDIRGRDGLKCDYNSKQKPKIKHMPRL
              SRMGNRSKIACHLPGRDNEIKYRWRTQKAKILKCDYNSKQKPKIKHMPRL
              IERTAAIOSVQFTSNHSPENSVATISTSSSEAVRSFYGDOVERGTDHMTNG
              GYWPNGGDTEFTLCSFDELKWLIO"
              816..931
              897
              3'UTR
              polyA_site
              /note="alternative"
BASE COUNT   317 a 183 c 194 g 237 t
ORIGIN
Query Match 17.3%; Score 149.2; DB 8; Length 931;
Best Local Similarity 67.2%; Pred. No. 3.7e-22;
Matches 211; Conservative 0; Mismatches 103; Indels 0; Gaps 0;

```

```

DB 414 AAGATTGCTCAACNTTTRCCGGGAAGACAGATACGAGTAAAGATTATTGGAGACA 473
OY 368 AGGATCCGAAGCA 381
DB 474 CGCGTTCAAAAGCA 487

RESULT 14
AF062894 1024 bp mRNA linear PLN 09-DEC-1998
LOCUS     Arabidopsis thaliana putative transcription factor (MYB59) mRNA,
DEFINITION complete cds.
ACCESSION AF062894
VERSION    AF062894.1 GI:3941479
KEYWORDS
SOURCE     Arabidopsis thaliana.
ORGANISM   Arabidopsis thaliana; Streptophyta; Embryophyta; Tracheophyta;
              Eukaryota; Viridiplantae; Magnoliophyta; eudicotyledons; core eudicots;
              Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1 (bases 1 to 1024)
AUTHORS     Kranz,H.D., Denekamp,M., Greco,R., Jin,H., Leyva,A., Meisner,R.C.,
              Petroni,K., Urzainqui,A., Beyan,M., Martin,C., Smeekens,S.,
              Tonelli,C., Paz-Ares,J. and Weisshaar,B.
              Towards functional characterisation of the members of the R2R3-MYB
              gene family from Arabidopsis thaliana
              Plant J. 16 (2), 263-276 (1998)
JOURNAL     9839469
MEDLINE     99056848
PUBMED      9839469
REFERENCE   2 (bases 1 to 1024)
AUTHORS     Kranz,H.D. and Weisshaar,B.
JOURNAL     Direct Submission
TITLE       Submitted (04-MAY-1998) Abt. Biochemie, MPI fur Zuechtungsforchung,
              Carl-von-Linne-Weg 10, Koeln D-50829, Germany
FEATURES
  source      1..1024
              location/Qualifiers
              /organism="Arabidopsis thaliana"
              /cultivar="Columbia"
              /db_xref="taxon:3702"
              1..1024
              /gene="MYB59"
              /allele="COL"
              70..774
              /gene="MYB59"
              /note="MYB59; R2R3-MYB factor family member"
              /codon_start=1
              /product="putative transcription factor"
              /protein_id="AAC83616.1"
              /db_xref="GI:3941480"
              /translation="MKLYOERYRKGRPTQEDLLVNFVHLGDRMDVVAVSGINR
              TKSCLRHWNVNLPGLKRGKMTPOERLVLELHAHWNKWSLIARKLGRDNEIKN
              YMRHMRKKADEKRRPMSPTSSNCCSSMTTQDYGSGNKGKNOGEGYVSM
              DIMREIDSGHIVKIPVDNYSEOSCLNPPPLASPTWSSIESIWNMDAESKMSF
              AIDQPLSFEGHSGL"
BASE COUNT   344 a 183 c 230 g 267 t
ORIGIN
Query Match 17.3%; Score 149.2; DB 8; Length 1024;
Best Local Similarity 67.2%; Pred. No. 3.6e-22;
Matches 211; Conservative 0; Mismatches 103; Indels 0; Gaps 0;

```



GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

## OM nucleic - nucleic search, using sw model

Run on: February 18, 2003, 08:10:00 ; Search time 48 Seconds  
(without alignments)  
5513.793 Million cell updates/sec

Title: US-10-021-811-35

Perfect score: 863  
Sequence: 1 gcacgagctctatcacacac.....aaaaaaaaaaaaaaaaaaaaa 863

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents\_NA:\*

- 1: /cgn2\_6/ptodata/1/lna/5A.COMB.seq:\*
- 2: /cgn2\_6/ptodata/1/lna/5B.COMB.seq:\*
- 3: /cgn2\_6/ptodata/1/lna/5A.COMB.seq:\*
- 4: /cgn2\_6/ptodata/1/lna/5B.COMB.seq:\*
- 5: /cgn2\_6/ptodata/1/lna/PCRTUS.COMB.seq:\*
- 6: /cgn2\_6/ptodata/1/lna/Backfiles1.seq:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	112.4	13.0	2220	4	US-08-997-251-1
2	110.8	12.8	2352	4	US-08-997-251-3
3	108.2	12.5	1344	2	US-08-722-626B-1
4	87	10.1	3602	4	US-09-402-923-1
5	87	10.1	6775	4	US-09-402-923-4
6	83.8	9.7	4880	4	US-09-402-923-5
7	78.8	9.1	1035	5	PCT-US93-06251-90
8	78.8	9.1	3225	5	US-08-306-691B-45
9	78.8	9.1	3225	5	PCT-US93-06251-91
10	78.8	9.1	3230	5	PCT-US93-06251-89
11	72.4	8.4	1897	4	US-09-167-322-10
12	60	7.0	2638	1	US-08-306-691B-46
13	52.8	6.1	7218	1	US-08-232-463-14
14	51.6	6.0	3573	4	US-09-353-585-4
15	51.2	5.9	1007	4	US-09-171-156A-61
16	51.2	5.9	1007	4	US-09-171-156A-63
17	51	5.9	1641	1	US-08-300-903A-8
18	51	5.9	4059	2	US-08-485-139-1
19	51	5.9	4059	3	US-08-485-139-1
20	51	5.9	4824	2	US-08-485-139-5
21	51	5.9	4824	3	US-08-485-139-5
22	50.4	5.8	2007	3	US-08-747-221B-36
23	50.4	5.8	2007	3	US-08-747-221B-38
24	50.4	5.8	2007	4	US-09-005-051-36
25	50.4	5.8	2007	4	US-09-005-051-38
26	49.4	5.7	3138	1	US-07-867-106-4
27	48.8	5.7	1582	3	US-08-545-196B-10

28	48.8	5.7	1582	3	US-08-545-196B-12	Sequence 12, Appl
29	48	5.6	1577	4	US-08-821-994-59	Sequence 59, Appl
30	47.8	5.5	1313	4	US-09-149-476-112	Sequence 112, App
31	47.8	5.5	1474	4	US-08-821-994-64	Sequence 64, Appl
32	47.6	5.5	785	3	US-09-008-979A-6	Sequence 6, Appl1
33	47.6	5.5	785	4	US-09-460-618-6	Sequence 6, Appl1
34	47.6	5.5	785	4	US-08-310-235B-6	Sequence 63, Appl
35	47.4	5.5	1441	4	US-08-821-994-63	Sequence 2, Appl1
36	47.4	5.5	5852	1	US-07-867-106-2	Sequence 53, Appl
37	47	5.4	370	4	US-09-376-113-1	Sequence 3, Appl1
38	45.8	5.3	1100	3	US-09-248-335-33	Sequence 56, Appl
39	45.8	5.3	1307	4	US-09-250-609-3	Sequence 1, Appl1
40	45.8	5.3	1753	4	US-09-149-476-56	Sequence 49, Appl
41	45.6	5.3	2908	4	US-09-930-181-1	Sequence 1, Appl1
42	45.4	5.3	1781	4	US-09-499-302A-1	Sequence 49, Appl
43	45.2	5.2	760	1	US-08-276-452A-49	Sequence 1, Appl1
44	45.2	5.2	760	2	US-08-798-744-49	Sequence 49, Appl
45	45.2	5.2	1889	3	US-09-187-050-1	Sequence 1, Appl1

## ALIGNMENTS

RESULT 1  
US-08-997-251-1  
; Sequence 1, Application US/08997251  
; Patent No. 6271440  
; GENERAL INFORMATION:  
; APPLICANT: GUBLER, FRANZ J.  
; APPLICANT: JACOBSEN, JOHN V.  
; TITLE OF INVENTION: PLANT REGULATORY PROTEINS III  
; NUMBER OF SEQUENCES: 17  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Greenlee, Winner and Sullivan, P.C.  
; STREET: 5370 Manhattan Boulevard  
; CITY: Boulder  
; STATE: CO  
; COUNTRY: US  
; ZIP: 80303  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/997,251  
; FILING DATE: 23-DEC-1997  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: WO AU96/00383  
; FILING DATE: 21-JUN-1996  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: AU PN6470/95  
; FILING DATE: 09-NOV-1995  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: AU PN3779/95  
; FILING DATE: 23-JUN-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Winner, Ellen P.  
; REGISTRATION NUMBER: 28,547  
; REFERENCE/DOCKET NUMBER: 110-97  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (303) 499-8080  
; TELEFAX: (303) 499-8089  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 2220 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA  
; ORGANISM: Hordeum vulgare





```

1  APPLICATION NUMBER: US/08/722,626B
2
3  FILING DATE: 27-SEP-1996
4
5  CLASSIFICATION: 800
6
7  PRIOR APPLICATION DATA:
8
9  APPLICATION NUMBER:
10
11  FILING DATE:
12
13  ATTORNEY/AGENT INFORMATION:
14
15  NAME: Pat Hagan
16
17  REGISTRATION NUMBER: 27,643
18
19  REFERENCE/DOCKET NUMBER: 97-0010
20
21  TELECOMMUNICATION INFORMATION:
22
23  TELEPHONE: 215 563-4100
24
25  TELEFAX: 215 563-4044
26
27  TELEX:
28
29  INFORMATION FOR SEQ ID NO: 1:
30
31  SEQUENCE CHARACTERISTICS:
32
33  LENGTH: 1344 base pairs
34
35  TYPE: nucleic acid
36
37  STRANDEDNESS: single
38
39  TOPOLOGY: linear
40
41  MOLECULE TYPE: cDNA
42
43  HYPOTHETICAL: NO
44
45  ANTI-SENSE: NO
46
47  FRAGMENT TYPE:
48
49  ORIGINAL SOURCE:
50
51  FEATURE:
52
53  NAME/KEY: Coding Sequence
54
55  LOCATION: 148..981
56
57  OTHER INFORMATION:
58
59  US-08-722-626B-1

```

```

1 ADDRESS: SEIDEL, GONDA, LAVOGNA & MONACO, P.C.
2 STREET: Suite 1800 Two Penn Center Plaza
3 CITY: Philadelphia
4 STATE: PA
5 COUNTRY: U.S.A.
6 ZIP: 19102
7
8 COMPUTER READABLE FORM:
9 MEDIUM TYPE: Floppy disk
10 COMPUTER: IBM PC compatible
11 OPERATING SYSTEM: PC-DOS/MS-DOS
12 SOFTWARE: PatentIn Release #1.0, Version #1.30
13
14 CURRENT APPLICATION DATA:
15 APPLICATION NUMBER: US/09/402,929
16 FILING DATE:
17
18 CLASSIFICATION:
19 PRIOR APPLICATION DATA:
20 APPLICATION NUMBER: PCT/US98/06896
21 FILING DATE:
22
23 ATTORNEY/AGENT INFORMATION:
24 NAME: Monaco, Daniel A.
25 REGISTRATION NUMBER: 30,480
26 REFERENCE/DOCKET NUMBER: 6056-214 PC
27
28 TELECOMMUNICATION INFORMATION:
29 TELEPHONE: (215) 568-8383
30 TELEFAX: (215) 568-5549
31
32 INFORMATION FOR SEQ ID NO: 1:
33 SEQUENCE CHARACTERISTICS:
34 LENGTH: 3602 base pairs
35 TYPE: nucleic acid
36 STRANDEDNESS: single
37 TOPOLOGY: linear
38
39 US-09-402-929-1

```



OY 399 ACTTTCA 405  
| | | | |  
Db 682 ATTACA 688

## RESULT 7

PCT-US93-06251-90  
; Sequence 90, Application PC/TUS9306251  
; GENERAL INFORMATION:  
; APPLICANT: Wickstrom, Eric and Rife, Jason P.  
; TITLE OF INVENTION: Trivalent Synthesis of Oligonucleotides Containing  
; TITLE OF INVENTION: Stereospecific Alkylphosphonates and Arylphosphonates  
; NUMBER OF SEQUENCES: 93  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: SCULLY, SCOTT, MURPHY & PRESSER  
; STREET: 400 Garden City Plaza  
; CITY: Garden City  
; STATE: NY  
; COUNTRY: USA  
; ZIP: 11530  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION NUMBER: PCT/US93/06251  
; FILING DATE: 19930630  
; CLASSIFICATION:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Digilio, Frank S.  
; REGISTRATION NUMBER: 31,346  
; REFERENCE/DOCKET NUMBER: 8586  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 516-742-4343  
; TELEFAX: 516-742-4366  
; TELEX: 230 901 SANS UR  
; INFORMATION FOR SEQ ID NO: 90:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1035 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (genomic)  
; PCT-US93-06251-90

Query Match 9.1%; Score 78.8; DB 5; Length 1035;  
Best Local Similarity 52.7%; Pred. No. 6.8e-10;

Matches 195; Conservative 0; Mismatches 172; Indels 3; Gaps 1;

OY 39 AACACAGCTTAAGACGCTCAAGATCCTGAAGTGAAGAAAGGCGCTTGACAATGGAAG 98  
| | | | |  
Db 98 AGACCGATGGCAGAAAGTACTAAACCTGACCTCATGAGGCTCTTGACCAAGAAG 157  
| | | | |  
OY 99 AAGACTTGATCTGATGAATATATATGCAATATATGGAAGTGTGGAATCTTTGG 158  
| | | | |  
Db 158 AAGATCAGAGAGTATGAGCTTGTACAGAAATACGCTCGAAGCTGTGCTGTTATTG 217  
| | | | |  
OY 159 CCAAGCTGCTGCTCAACAGTACGGAAGAGTTCGCGCTAAGTGGCTAAATTACC 218  
| | | | |  
Db 218 CCAAGCACTTAAGGGA--GAATTGGAAGAAACATGTAGGAGAGTGGCATTAACCACT 274  
| | | | |  
OY 219 TCCGTCCTGATTTAGAAGAGGAATATTACACCCGAGAACCACTTTGATTATGAGAC 278  
| | | | |  
Db 275 TGAATCCAGAAAGTTAAGAAACCTCCTGGCAGAAAGAGACAGAAATATTATACCAAG 334  
| | | | |  
OY 279 TCCACCAAGAGTGGGAACAGAGTGTCCAAATTTGCCAAGCATCTACCTGGAAGACAG 338  
| | | | |  
Db 335 CACACAGAGAGCTGGGGAACAGATGGGCAAGAAATCCCAAGCTACTGCTGGACGAGACAG 394  
| | | | |  
OY 339 ATATATGATCAAGAACTATTTGAGAGCAAGATCCAGAGACATCAAGCAAGCTGACA 398  
| | | | |  
Db 395 ATATATGATCAAGAACTATTTGAGAGCAAGATCCAGAGACATCAAGCAAGCTGACA 454  
| | | | |

OY 399 ACTTTCAGCA 408  
| | | | |  
Db 455 ATCTGCAGCA 464

## RESULT 8

US-08-306-691B-45  
; Sequence 45, Application US/0830691B  
; Patent No. 5734039  
; GENERAL INFORMATION:  
; APPLICANT: Calabretta, Bruno  
; APPLICANT: Skorski, Tomasz  
; TITLE OF INVENTION: ANTISENSE  
; TITLE OF INVENTION: OLIGONUCLEOTIDES TARGETING COOPERATING ONCOGENES  
; NUMBER OF SEQUENCES: 55  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Seidel, Gonda, Lavorgna & Monaco, P.C.  
; STREET: Two Penn Center, Suite 1800  
; CITY: Philadelphia  
; STATE: Pennsylvania  
; COUNTRY: U.S.A.  
; ZIP: 19102  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.50 inch, 720 Kb  
; COMPUTER: IBM PS/2  
; OPERATING SYSTEM: MS-DOS  
; SOFTWARE: Wordperfect 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/306,691B  
; FILING DATE: September 15, 1994  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Monaco, Daniel A.  
; REGISTRATION NUMBER: 30,480  
; REFERENCE/DOCKET NUMBER: 8321-8  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (215) 568-8383  
; TELEFAX: (215) 568-5549  
; TELEX: No. 5734039e  
; INFORMATION FOR SEQ ID NO: 45:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 3225 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
; US-08-306-691B-45

Query Match 9.1%; Score 78.8; DB 1; Length 3225;  
Best Local Similarity 52.7%; Pred. No. 8.8e-10;

Matches 195; Conservative 0; Mismatches 172; Indels 3; Gaps 1;

OY 39 AACACAGTGTAAAGCGTCTCAAGATCCTGAAGTGAAGAAAGGCGCTTGACAATGGAAG 98  
| | | | |  
Db 349 AGACCGATGGCAGAAAGTACTAAACCTGACCTCATGAGGCTCTTGACCAAGAAG 408  
| | | | |  
OY 99 AAGACTTGATCTGATGAATATATATGCAATATATGGAAGTGTGGAATCTTTGG 158  
| | | | |  
Db 409 AAGATCAGAGAGTATGAGCTTGTACAGAAATACGCTCGAAGCTGTGCTGTTATTG 468  
| | | | |  
OY 159 CCAAGCTGCTGCTCAACAGTACGGAAGAGTTCGCGCTAAGTGGCTAAATTACC 218  
| | | | |  
Db 469 CCAAGCACTTAAGGGA--GAATTGGAAGAAACATGTAGGAGAGTGGCATTAACCACT 525  
| | | | |  
OY 219 TCCGTCCTGATTTAGAAGAGGAATATTACACCCGAGAGCAACTTTGATTATGAGAC 278  
| | | | |  
Db 526 TGAATCCAGAAAGTTAAGAAACCTCCTGGCAGAAAGAGAGACAGATTTATTACCAAG 585  
| | | | |  
OY 279 TCCACCAAGAGTGGGAACAGAGTGTCCAAATTTGCCAAGCATCTACCTGGAAGACAG 338  
| | | | |

QY	Db	QY	Db
339	586	339	586
ATATGAGTATCAAAACATTGTGGAGCAAAAGATCCGAAAGCAGCATCAAGCAAGCTGAGA	CACACAAAGAGACTGGGGGAACGATATGGGCAGAGAAATCCGAAAAGCTACTCCCTGGACAGAACTG	ATATGAGTATCAAAACATTGTGGAGCAAAAGATCCGAAAGCAGCATCAAGCAAGCTGAGA	CACACAAAGAGACTGGGGGAACGATATGGGCAGAGAAATCCGAAAAGCTACTCCCTGGACAGAACTG
646	706	646	706
ATAATGCTATCAAGAACACCTGGATTTCTACAAATGCGTCGGAAGGTGGAAACAGAAAGTT	ATCTGCAGAGCACTGGATTTCTACAAATGCGTCGGAAGGTGGAAACAGAAAGTT	ATAATGCTATCAAGAACACCTGGATTTCTACAAATGCGTCGGAAGGTGGAAACAGAAAGTT	ATCTGCAGAGCACTGGATTTCTACAAATGCGTCGGAAGGTGGAAACAGAAAGTT
QY	Db	QY	Db
339	586	339	586
ACTTTCAGCA	ACTTTCAGCA	ACTTTCAGCA	ACTTTCAGCA
408	408	408	408
QY	Db	QY	Db
706	706	706	706
ATCTGCAGCA	ATCTGCAGCA	ATCTGCAGCA	ATCTGCAGCA
715	715	715	715

```

1      RESULT 9
2      PCT-US93-06251-91
3      Sequence 91, Application PC/US9306251
4      GENERAL INFORMATION:
5      APPLICANT: WICKSTROM, Eric and Rife, Jason P.
6      TITLE OF INVENTION: Trivalent Synthesis of Oligonucleotides Containing
7      TITLE OF INVENTION: Stereospecific Alkylphosphonates and Arylphosphonates
8      NUMBER OF SEQUENCES: 93
9      CORRESPONDENCE ADDRESS:
10     ADDRESSEE: SCULLY, SCOTT, MURPHY & PRESSER
11     STREET: 400 Garden City Plaza
12     CITY: Garden City
13     STATE: NY
14     COUNTRY: USA
15     ZIP: 11530
16     COMPUTER READABLE FORM:
17     MEDIUM TYPE: Floppy disk
18     COMPUTER: IBM PC compatible
19     OPERATING SYSTEM: PC-DOS/MS-DOS
20     SOFTWARE: PatentIn Release #1.0, Version #1.25
21     CURRENT APPLICATION DATA:
22     APPLICATION NUMBER: PCT/US93/06251
23     FILING DATE: 19930630
24     CLASSIFICATION:
25     ATTORNEY/AGENT INFORMATION:
26     NAME: Digiglio, Frank S.
27     REGISTRATION NUMBER: 31,346
28     REFERENCE/DOCKET NUMBER: 8586
29     TELECOMMUNICATION INFORMATION:
30     TELEPHONE: 516-742-4343
31     TELEFAX: 516-742-4366
32     TELEX: 230 901 SANS UR
33     INFORMATION FOR SEQ ID NO: 91:
34     SEQUENCE CHARACTERISTICS:
35     LENGTH: 3225 base pairs
36     TYPE: nucleic acid
37     STRANDEDNESS: double
38     TOPOLOGY: linear
39     MOLECULE TYPE: DNA (genomic)
40     PCT-US93-06251-91

```

[illegible]

Accession	Sequence	Position
Db	586 CACACAAAGAGACTGGGGAAACGATGGGGAGAAATCGCAAAAGCTACTGCTGGAGCAAACTG	645
OY	339 ATATGAGATTAAGAACTATTGGAGGACAAAGATCCAGAAAGCACATTAAGCAAGCTGAGA	398
Db	646 ATAAATGCTATCAAGAACACACGTGAATTTCACTAAATGGCGTTCGGAAAGCTGAACAGAGAAAGTT	705
OY	399 ACTTTCAGCA 408	
Db	706 ATCTGCAGGA 715	

```

10 RESULT
11 PCT-US93-06251-89
12 : Sequence 89, Application PC/TUS9306251
13 : GENERAL INFORMATION:
14 : APPLICANT: WICKSTROM, Eric and Rife, Jason P.
15 : TITLE OF INVENTION: Trivalent Synthesis of Oligonucleotides Containing
16 : TITLE OF INVENTION: Stereospecific Alkylphosphonates and Arylphosphonates
17 : NUMBER OF SEQUENCES: 93
18 : CORRESPONDENCE ADDRESS:
19 : ADDRESSEE: SCULLY, SCOTT, MURPHY & PRESSER
20 : STREET: 400 Garden City Plaza
21 : CITY: Garden City
22 : STATE: NY
23 : COUNTRY: USA
24 : ZIP: 11530
25 : COMPUTER READABLE FORM:
26 : MEDIUM TYPE: Floppy disk
27 : COMPUTER: IBM PC compatible
28 : OPERATING SYSTEM: PC-DOS/MS-DOS
29 : SOFTWARE: PatentIn Release #1.0, Version #1.25
30 : CURRENT APPLICATION DATA:
31 : APPLICATION NUMBER: PCT/US93/06251
32 : FILING DATE: 19930630
33 : CLASSIFICATION:
34 : ATTORNEY/AGENT INFORMATION:
35 : NAME: DIGILIO, Frank S.
36 : REGISTRATION NUMBER: 31,346
37 : REFERENCE/DOCKET NUMBER: 8586
38 : TELECOMMUNICATION INFORMATION:
39 : TELEPHONE: 516-742-4343
40 : TELEFAX: 516-742-4366
41 : TELEX: 230 901 SANS UR
42 : INFORMATION FOR SEQ ID NO: 89:
43 : SEQUENCE CHARACTERISTICS:
44 : LENGTH: 3230 base pairs
45 : TYPE: nucleic acid
46 : STRANDEDNESS: double
47 : TOPOLOGY: linear
48 : MOLECULE TYPE: DNA (genomic)
49 : PCT-US93-06251-89

```

Query Match	Match Similarity	Score	DB	Length
Best Local	Similarity 52.7%	Pred. No. 8	8-10	
Matches	195	Conservative	0	Mismatches 172; Indels 3; Gaps 1;

  

QY	39	AACACAGTGAAGCGTCTCAAGATCCGAGAGTGAAGAAAGGGCTTGGCAATGTGAAG	98
Db	239	ACGACCGATGCGAGAAAGACTAAACCTGACCTCATCAAGSGTCTCTTGACAAAGAAG	298
QY	99	AAGCTTTGATCTTTGATGAACTATTTGCCAAATCATGCGGAAGCTGTTTGGAACTCTTGG	158
Db	299	AGATGACGAGAGTGAATGAGGCTTGACAGAAATACGGTCCGAACGTTGCTGTATTGG	358
QY	159	CCAAAGCTCTGCTCTCAAAAGCTAAACGGAAGAAGTTGCCCGCTAAAGTGCGCTAAATTACC	218
Db	359	CCAAAGCACTTAAAGGGGA---GAATTTGAAAAACAATGTAGGGAAGGTGCAATTAACACT	415
QY	219	TCCGTCCTGATGTTAGAAAGAGGGAATATTACACCCGAGGAGAACACTTTTGATTATGAGAC	278
Db	416	TCAATTCAGAACTTTAAGAAACCTCTCTGACGACGAAGAGAGAACAGCAAGATATTATTACCAAG	475





```

; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: CDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Tomato
; STRAIN: CR2
; SEQUENCE DESCRIPTION: SEQ ID NO: 4 :
US-09-353-585-4

Query Match
Best Local Similarity 51.8%; Pred. No. 0.0021;
Matches 114; Conservative 1; Mismatches 105; Indels 0; Gaps 0;

Qy 644 GATTAAATATATATCAAGATTAACCTAAGTTTGAAGTTCATACGCTGGAATGCTCTT 703
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3352 GATTGATTTCAAGACTTCGACTTCAGAGCCAGAAATAGAAAGCGTGTGTAAG 3411
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 704 GGAATTAACATATTTATGCGTTTGTATATAGTAGTTGATGTTTGGCTTAC 763
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3412 GATTGCTTCTTCCTGCTGCTGACGCTTATGATGTTGATGATGATTTAGTTTATAAGC 3471
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 764 CATTATAGCTATGCTGTAAATATATACGAGATTTATTTAACTATATCTGCATGCT 823
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3472 TTTTCTTCACTTGAGAAAGTATATATTTATGAATTTGATGATTAACATTAAGTGTGTG 3531
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 824 TTATATATATAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 863
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3532 TTATTGAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3571
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 15
US-09-171-156A-61
; Sequence 61, Application US/09171156A
; Patent No. 6368846
; GENERAL INFORMATION:
; APPLICANT: Hunter, Shirley Wu
; Sm, Gek-Ke
; Weber, Eric R.
; TITLE OF INVENTION: NOVEL ECTOPARASITE SALIVA PROTEINS AND
; APPARATUS TO COLLECT SUCH PROTEINS
; NUMBER OF SEQUENCES: 88
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SHERIDAN ROSS P.C.
; STREET: 1560 BROADWAY, SUITE 1200
; CITY: DENVER
; STATE: CO
; COUNTRY: U.S.A.
; ZIP: 80202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/171,156A
; FILING DATE: 04-Mar-1999
; CLASSIFICATION: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Conneil, Gary J.
; REGISTRATION NUMBER: 32,020
; REFERENCE/DOCKET NUMBER: 2618-17-C4-PUS
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 303/863-9700
; TELEFAX: 303/863-0223
; INFORMATION FOR SEQ ID NO: 61:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1007 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
```

```

; MOLECULE TYPE: CDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..465
; SEQUENCE DESCRIPTION: SEQ ID NO: 61 :
US-09-171-156A-61

Query Match
Best Local Similarity 5.9%; Score 51.2; DB 4; Length 1007;
Matches 95; Conservative 1; Mismatches 74; Indels 0; Gaps 0;

Qy 692 TGAATGCTCTTGGAATTAACATATTTATGCGTTGTTTATATAGAGTGGATGTT 751
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 838 TGGGGATGTTATATATATAAAGTAGTGATTTATTTATTAAGTTCAGAAA 897
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 752 GCTTTGCGTACCATTTATAGCTATGCTGCTGAATATATACGAGATTTATATTAAC 811
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 898 TGTTTCATCATTAATTAACGAAATTTTCAATATATATTTGATGATTAATCGC 957
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 812 TATCTGCATGCTTTATATATAAAAAAAAAAAAAAAAAAAAAA 861
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 958 CATCTGATGCTTAATATTAATAAAAAAAAAAAAAAAAAAAAAA 1007
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

Search completed: February 18, 2003, 09:20:41  
Job time : 71 secs





GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 18, 2003, 07:07:50 ; Search time 218 Seconds

(without alignments)  
8915.021 Million cell updates/sec

Title: US-10-021-811-35

Perfect score: 863  
Sequence: 1 gcaacgagcctatcaacac.....aaaaaaaaaaaaaaaaa 863

Scoring table:  
IDENTITY\_NUC  
Gapop 10.0 , Gapept 1.0

Searched: 2185239 segs, 112599159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :

N\_Geneseq\_101002: \*  
1: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT: \*  
2: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT: \*  
3: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT: \*  
4: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT: \*  
5: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT: \*  
6: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT: \*  
7: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT: \*  
8: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT: \*  
9: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT: \*  
10: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT: \*  
11: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT: \*  
12: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT: \*  
13: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT: \*  
14: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT: \*  
15: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT: \*  
16: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT: \*  
17: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT: \*  
18: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT: \*  
19: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT: \*  
20: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT: \*  
21: /SID22/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT: \*  
22: /SID22/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT: \*  
23: /SID22/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT: \*  
24: /SID22/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	248.4	28.8	862	21	AAC57336
2	239.4	27.7	910	21	AAC40587
3	236.2	27.4	908	21	AAC48506
4	223	25.8	907	21	AAC57337
5	208	24.1	621	21	AAC41714
6	208	24.1	1127	21	AAC37953
7	175	20.3	626	22	AAH87724
8	170.6	19.8	1137	22	AAD05784
9	169.4	19.6	837	20	AAK25572

10	169.4	19.6	981	24	ABK65252	Arabidopsis cDNA e
11	165.2	19.1	673	21	AAC56197	Eucalyptus grandis
12	163.4	18.9	972	24	ABK65364	Arabidopsis cDNA e
13	161.6	18.7	389	21	AAC56152	Eucalyptus grandis
14	161.6	18.7	417	21	AAC56774	Eucalyptus grandis
15	152.4	17.7	524	21	AAC57194	Eucalyptus grandis
16	143.2	16.6	959	21	AAC44894	Arabidopsis thalia
17	132.8	15.4	516	21	AAC56352	Pinus radiata tran
18	132.8	15.4	516	21	AAC56457	Pinus radiata tran
19	131.4	15.2	542	21	AAC56434	Pinus radiata tran
20	130.8	15.2	839	21	AAC42479	Pinus radiata tran
21	130.8	15.2	839	21	AAC42479	Arabidopsis thalia
22	130.6	15.1	918	22	AAD06466	Arabidopsis thalia
23	129.6	15.0	389	21	AAC56837	Pinus radiata tran
24	128.6	14.9	984	21	AAC44207	Arabidopsis thalia
25	128.4	14.9	1329	21	AAC49321	Arabidopsis thalia
26	128	14.8	1329	21	AAC49321	Arabidopsis thalia
27	127	14.7	1205	21	AAC47877	Arabidopsis thalia
28	126.8	14.7	1252	24	ABK65304	Arabidopsis thalia
29	126.8	14.7	1332	21	AAC40461	Arabidopsis thalia
30	125.4	14.5	734	21	AAC54477	Arabidopsis thalia
31	125.4	14.5	736	21	AAC52742	Arabidopsis thalia
32	125.2	14.5	612	24	ABK65317	Arabidopsis thalia
33	124.4	14.4	1152	21	AAC46330	Arabidopsis cDNA e
34	124.4	14.4	1153	21	AAC39525	Arabidopsis thalia
35	123.2	14.3	373	21	AAC57227	Eucalyptus grandis
36	123	14.3	378	21	AAC56151	Eucalyptus grandis
37	122.8	14.2	639	21	AAC51202	Arabidopsis thalia
38	122.8	14.2	916	22	AAD05774	Arabidopsis thalia
39	122.6	14.2	825	21	AAC42974	Arabidopsis thalia
40	122.6	14.2	825	22	AAD06465	Arabidopsis thalia
41	121.6	14.1	1151	22	AAF90595	Cotton transcript
42	121.6	14.1	1300	24	ABK65274	Arabidopsis thalia
43	120.8	14.0	774	21	AAC42662	Arabidopsis thalia
44	120.8	14.0	1044	24	ABN98399	Arabidopsis thalia
45	120.6	14.0	1678	21	AAC57334	Eucalyptus grandis

#### ALIGNMENTS

RESULT 1  
ID AAC57336 standard; DNA: 862 BP.  
AAC57336;  
25-JAN-2001 (first entry)  
Eucalyptus grandis transcription factor DNA sequence #773.  
Plant; transcription factor; gene expression; eucalyptus; pine; acacia;  
poplar; sweetgum; teak; mahogany; bZIP; G-box binding factor;  
basic helix-loop-helix zipper; homeotic; homeobox; MADS;  
homeodomain zipper; LIM domain; AP2; ERBBS; zinc finger domain;  
type 2 Cys2His2; CCAAT box element; MYB; ss.  
Eucalyptus grandis.  
WO200053724-A2.  
14-SEP-2000.  
09-MAR-2000; 2000WO-US06112.  
11-MAR-1999; 99US-0266513.  
18-AUG-1999; 99US-0149485.  
(GENE-) GENESIS RES & DEV CORP LTD.  
(FLET-) FLETCHER CHALLENGE FORESTS LTD.  
Wood M. McGrath A. Shank MA. Glenn M.



PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 02-JUL-1999; 99US-0142154.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142970.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145813.  
PR 27-JUL-1999; 99US-0145819.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 05-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 06-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147203.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 11-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 23-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.  
PR 25-AUG-1999; 99US-0149930.  
PR 26-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151080.

PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158332.  
PR 12-OCT-1999; 99US-0158369.  
PR 13-OCT-1999; 99US-0159293.  
PR 13-OCT-1999; 99US-0159294.  
PR 13-OCT-1999; 99US-0159295.  
PR 14-OCT-1999; 99US-0159329.  
PR 14-OCT-1999; 99US-0159330.  
PR 14-OCT-1999; 99US-0159331.  
PR 14-OCT-1999; 99US-0159637.  
PR 14-OCT-1999; 99US-0159638.  
PR 18-OCT-1999; 99US-0159854.  
PR 21-OCT-1999; 99US-0160741.  
PR 21-OCT-1999; 99US-0160767.  
PR 21-OCT-1999; 99US-0160768.  
PR 21-OCT-1999; 99US-0160770.  
PR 21-OCT-1999; 99US-0160814.  
PR 21-OCT-1999; 99US-0160815.  
PR 22-OCT-1999; 99US-0160980.  
PR 22-OCT-1999; 99US-0160981.  
PR 22-OCT-1999; 99US-0160989.  
PR 25-OCT-1999; 99US-0161404.  
PR 25-OCT-1999; 99US-0161405.  
PR 25-OCT-1999; 99US-0161406.  
PR 26-OCT-1999; 99US-0161359.  
PR 26-OCT-1999; 99US-0161360.  
PR 26-OCT-1999; 99US-0161361.  
PR 28-OCT-1999; 99US-0161920.  
PR 28-OCT-1999; 99US-0161922.  
PR 28-OCT-1999; 99US-0161993.  
PR 29-OCT-1999; 99US-0162142.

Query Match 27.7%; Score 239.4; DB 21; Length 910;  
Best Local Similarity 72.4%; Pred. No. 1e-38;  
Matches 325; Conservative 0; Mismatches 121; Indels 3; Gaps 1;  
QY 56 TCTCAAGATCCGAGAGTGAAGAAAGGCCCTTGACATGGAAGAGACTGATCTTGATG 115  
DB 143 TCAGAGATGACAGAGTGAAGAAAGGCCATGACGATGGAAGAGATTATCTCATC 202  
QY 116 AACATATTGCAAAATCATGGGGAAGGTGTTGGAACTCTTGGCCAAAGCTCGTGC 175  
DB 203 AATTATTCGCCAATCATCATGTTGAGGTCTTTGAACTCTTCGCCAAATCTGCAGAGCTA 262  
QY 176 AAACGTAAACGGAAGAGTTGCCGCGCTAAGTGGCTAAATTACCTCCGCTCGATGTTAGA 235  
DB 263 AAACGCACCGGGAAGAAAGTTGCCGCGCTCGGCTGGAACCTACCTCCGACCTGATGTCGA 322  
QY 236 AGAGGAAATATTACACCCGAGGAACAACTTTGATTTAGAGCTTCACGCAAGTGGGA 295  
DB 323 CGGGAAATATCACACCAAGAACAGCTCACATCATGAACTTCAATGCAAAATGGGGA 382  
QY 296 AACAGTGTCCAAAAATTTGCCAAGCATCTACCTGAAGAGCTGAATTAATGAGATCAAGAC 355  
DB 383 AATAGTGTGTCAAAATTTGCAAGAGATTTCACGAGGAAGACCGACATGAGATTAAGAAC 442

PR	07-JUN-1999	9905-0137724
PR	08-JUN-1999	9905-0138094
PR	10-JUN-1999	9905-0138540
PR	10-JUN-1999	9905-0138847
PR	14-JUN-1999	9905-0139119
PR	16-JUN-1999	9905-0139452
PR	16-JUN-1999	9905-0139453
PR	17-JUN-1999	9905-0139492
PR	17-JUN-1999	9905-0139454
PR	18-JUN-1999	9905-0139455
PR	18-JUN-1999	9905-0139456
PR	18-JUN-1999	9905-0139462
PR	18-JUN-1999	9905-0139463
PR	18-JUN-1999	9905-0139750
PR	18-JUN-1999	9905-0139763
PR	21-JUN-1999	9905-0139817
PR	22-JUN-1999	9905-0139899
PR	23-JUN-1999	9905-0140353
PR	23-JUN-1999	9905-0140354
PR	24-JUN-1999	9905-0140695
PR	28-JUN-1999	9905-0140823
PR	29-JUN-1999	9905-0140991
PR	30-JUN-1999	9905-0141287
PR	01-JUL-1999	9905-0141842
PR	01-JUL-1999	9905-0142154
PR	02-JUL-1999	9905-0142050
PR	06-JUL-1999	9905-0142390
PR	08-JUL-1999	9905-0142803
PR	09-JUL-1999	9905-0142920
PR	12-JUL-1999	9905-0142977
PR	13-JUL-1999	9905-0143542
PR	14-JUL-1999	9905-0143624
PR	15-JUL-1999	9905-0144005
PR	16-JUL-1999	9905-0144085
PR	16-JUL-1999	9905-0144086
PR	19-JUL-1999	9905-0144325
PR	19-JUL-1999	9905-0144331
PR	19-JUL-1999	9905-0144332
PR	19-JUL-1999	9905-0144333
PR	19-JUL-1999	9905-0144334
PR	19-JUL-1999	9905-0144335
PR	20-JUL-1999	9905-0144352
PR	20-JUL-1999	9905-0144632
PR	21-JUL-1999	9905-0144884
PR	21-JUL-1999	9905-0144814
PR	21-JUL-1999	9905-0145086
PR	22-JUL-1999	9905-0145085
PR	22-JUL-1999	9905-0145087
PR	22-JUL-1999	9905-0145089
PR	22-JUL-1999	9905-0145192
PR	23-JUL-1999	9905-0145119
PR	23-JUL-1999	9905-0145218
PR	23-JUL-1999	9905-0145224
PR	26-JUL-1999	9905-0145276
PR	26-JUL-1999	9905-0145913
PR	27-JUL-1999	9905-0145918
PR	27-JUL-1999	9905-0145951
PR	28-JUL-1999	9905-0145951
PR	02-AUG-1999	9905-0146386
PR	02-AUG-1999	9905-0146388
PR	03-AUG-1999	9905-0146389
PR	03-AUG-1999	9905-0147038
PR	04-AUG-1999	9905-0147204
PR	04-AUG-1999	9905-0147302
PR	05-AUG-1999	9905-0147192
PR	05-AUG-1999	9905-0147260
PR	06-AUG-1999	9905-0147303

PR	06-AUG-1999;	99US-0147416.
PR	09-AUG-1999;	99US-0147493.
PR	09-AUG-1999;	99US-0147935.
PR	10-AUG-1999;	99US-0148171.
PR	11-AUG-1999;	99US-0148319.
PR	12-AUG-1999;	99US-0148341.
PR	13-AUG-1999;	99US-0148565.
PR	13-AUG-1999;	99US-0148684.
PR	16-AUG-1999;	99US-0149368.
PR	17-AUG-1999;	99US-0149175.
PR	18-AUG-1999;	99US-0149426.
PR	20-AUG-1999;	99US-0149722.
PR	20-AUG-1999;	99US-0149723.
PR	20-AUG-1999;	99US-0149929.
PR	23-AUG-1999;	99US-0149902.
PR	23-AUG-1999;	99US-0149930.
PR	23-AUG-1999;	99US-0150566.
PR	26-AUG-1999;	99US-0150884.
PR	27-AUG-1999;	99US-0151065.
PR	27-AUG-1999;	99US-0151066.
PR	27-AUG-1999;	99US-0151080.
PR	30-AUG-1999;	99US-0151303.
PR	31-AUG-1999;	99US-0151338.
PR	01-SEP-1999;	99US-0151930.
PR	07-SEP-1999;	99US-0152363.
PR	10-SEP-1999;	99US-0153070.
PR	13-SEP-1999;	99US-0153758.
PR	15-SEP-1999;	99US-0154018.
PR	16-SEP-1999;	99US-0154039.
PR	20-SEP-1999;	99US-0154779.
PR	22-SEP-1999;	99US-0155139.
PR	23-SEP-1999;	99US-0155486.
PR	24-SEP-1999;	99US-0155569.
PR	28-SEP-1999;	99US-0156458.
PR	29-SEP-1999;	99US-0156595.
PR	04-OCT-1999;	99US-0157117.
PR	05-OCT-1999;	99US-0157753.
PR	06-OCT-1999;	99US-0157865.
PR	07-OCT-1999;	99US-0158029.
PR	08-OCT-1999;	99US-0158232.
PR	12-OCT-1999;	99US-0158366.
PR	13-OCT-1999;	99US-0158293.
PR	13-OCT-1999;	99US-0159294.
PR	13-OCT-1999;	99US-0159295.
PR	14-OCT-1999;	99US-0159329.
PR	14-OCT-1999;	99US-0159330.
PR	14-OCT-1999;	99US-0159331.
PR	14-OCT-1999;	99US-0159637.
PR	14-OCT-1999;	99US-0159638.
PR	18-OCT-1999;	99US-0159584.
PR	21-OCT-1999;	99US-0160741.
PR	21-OCT-1999;	99US-0160767.
PR	21-OCT-1999;	99US-0160768.
PR	21-OCT-1999;	99US-0160770.
PR	21-OCT-1999;	99US-0160814.
PR	21-OCT-1999;	99US-0160815.
PR	22-OCT-1999;	99US-0160980.
PR	22-OCT-1999;	99US-0160981.
PR	22-OCT-1999;	99US-0160989.
PR	25-OCT-1999;	99US-0161404.
PR	25-OCT-1999;	99US-0161405.
PR	25-OCT-1999;	99US-0161406.
PR	26-OCT-1999;	99US-0161359.
PR	26-OCT-1999;	99US-0161360.
PR	26-OCT-1999;	99US-0161361.
PR	28-OCT-1999;	99US-0161920.
PR	28-OCT-1999;	99US-0161992.
PR	28-OCT-1999;	99US-0161993.
PR	29-OCT-1999;	99US-0162142.

OY	56	TCTCAAGATTCCTGGAAGTGACAAAAAGGCGCTTGAGACAAATGGAAGAAAGACTTGCATCTGATG	115		
Db	141	TCAGGAGATGCAAGAGGTGACAAAAAGGGCCATGACGATGGAAGAAAGATTTATTTCTCATC	200		
OY	116	AACATATATTCGAATTCATGGGGAAGGTGTTTGGAACTCTTTGGCCCAAAAGCTGCTGCTC	175		
Db	201	AATTATATTCGCCAATTCATGCTGTAAGGTGTTTGGAACTCTCTCCCAATTCGCAGACTA	260		
OY	176	AAACGTAACGGGAAGAGTGTCCGGGCTAAAGGTGCTAAATTTACCTCCGCTCCTGATGTTAGA	235		
Db	261	AAACGCCACCGGGAAAAATTTCCCGGCTCCG6TG6CTGAACCTACCTCCGACCTGATGTGCGA	320		
OY	236	AGAGGAGATATTTACACCCGAGCAACCTTTTGTATTATTCAGAGCTTCCAGCAAGTGGGGA	295		
Db	321	CGGGGAATATTCACACCAAGAAAGACGTCACACATCATGAAACTTCATGCAAAATGGGA	380		
OY	296	AAACAGGTGCTCCAAAATTTGGCCAGCATCTTACTCGGAGAGCATGTATGAGATCAGAAC	355		
Db	381	AATAGGTGGTCAAAAATTTGCAAAAGCATTTTACAGGAAAGACCGCAATGAATTAAGAT	440		
OY	356	TATTGGAGGACAMGATGTCGAAGCATCATCAGCAAGCTGAGAACTTTCAGCAACAGAT	415		
Db	441	TTTTGGAGGACTAAGATTCGAAATATTCATCATCAAGACCGGAGAAAGACAGCGTTGGA	500		
OY	416	AGTATATTTCTGTGATTAATGTATCCAAAGCTAGCACCTAGCCATGTTTCCACCATGGCT	475		
Db	501	TCACAAAGCTCCGAGTTTATTAACCATCATGTGACACAGACGACATGTCATGATGATACT	560		
OY	476	--GAGCCCATGAGATGATTTCTCCACC	501		
Db	561	CAAGAAACCATGATGATGTTCTCCAAC	589		
RESULT 4					
AC	AAC57337				
ID	AAC57337	standard; DNA; 907 BP.			
XX					
AC	AAC57337:				
XX					
DT	25-JAN-2001	(first entry)			
XX					
DE	Eucalyptus grandis transcription factor DNA sequence #774.				
XX					
KM	Plant; transcription factor; gene expression; eucalyptus; pine; acacia.				
KM	popular; sweetgum; leek; mahogany; bZIP; G-box binding factor;				
KM	basic helix-loop-helix zipper; homeotic; homeodomain; homeobox; MADS;				
KM	homeodomain zipper; LIM domain; AP2; ERBBs; zinc finger domain;				
XX	type 2 Cys2His2; COAT box element; MYB; ss.				
OS	Eucalyptus grandis.				
XX					
PN	MO200053724-A2.				
PD	14-SEP-2000.				
XX					
PF	09-MAR-2000;	2000MO-US06112.			
XX					
PR	11-MAR-1999;	99US-026613.			
PR	18-AUG-1999;	99US-0149485.			
XX					
PA	(GENE-) GENESIS RES & DEV CORP LTD.				
PA	(FLET-) FLETCHER CHALLENGE FORESTS LTD.				
XX					
PI	Wood M, McGrath A, Shank MA, Glenn M;				
XX					
DR	WPI; 2000-579369/54.				
XX					
PT	New isolated polynucleotide encoding a plant transcription factor for				
PT	producing a plant e.g. a woody plant, preferably eucalyptus or pine,				
PT	having modified gene expression or modified activity of a polypeptide				
XX					

PS Claim 1; Pages 649-650; 747pp; English.

XX The present invention relates to novel plant transcription factors from  
CC Eucalyptus grandis or Pinus radiata. The present sequence is the coding  
CC sequence for one such transcription factor. The transcription factor may  
CC be used to produce a plant having modified gene expression such as a  
CC woody plant e.g. a eucalyptus, pine, acacia, poplar, sweetgum, teak, or  
CC mahogany species or to modify the activity of a polypeptide in a plant.  
CC The transcription factors of the present invention are members from the  
CC following families of regulatory proteins: bZIP, bZIP family of G-box  
CC binding factors, basic helix-loop-helix zipper,  
CC homeotic/homeodomain/homeobox/MADS, homeodomain zipper, LIM domain, AP2  
CC and PEBBs, zinc finger domains of type 2 Cys2His2, CCAAT box elements  
CC and MYB.

XX Sequence 907 BP; 280 A; 210 C; 212 G; 205 T; 0 other;

Query Match 25.8%; Score 223; DB 21; Length 907;

Best Local Similarity 77.2%; Pred. No. 1.9e-35; Mismatches 271; Conservative 0; Indels 0; Gaps 0;

QY 36 AAAAAACAGTGAAGCGTCAAGATCTGAGTGAAGAAAGGCGCTTGACATGG 95  
DB 8 ACAAAGAAACCGTCTACCGCAGCAGATCTCTCAGTGAAGAAAGCAGCATGCTCTTG 67  
QY 96 AAGAAGACTTGATCTGTATGAACATATATGCAAAATCATGGGGAAGGTGTTGAACTCTT 155  
DB 68 AAGAGATCTGATCTGATGATGATACATAGCAACATGCGGGAAGGTATGAACTCTC 127  
QY 156 TGGCCAAAGCGTGTGCTCAAGCTTAAGGAAAGTGGCGCTTAAGGCGCTTAATT 215  
DB 128 TAGCTAAAGCTGTGCTGTCTTCAACGCAACGAAAGAGCTCGCGCTCGATGCTGAATT 187  
QY 216 ACTTCCTGCTGATGTAGAGAGGAATATTACACCCGAGAGCAACTTTGATTATGG 275  
DB 188 ACTTCGCGACCGATGTCGAGAGGTATATCACTCCGGAAGACACTGTGATCATAC 247  
QY 276 AGCTCCACGCAAGTGGGGAAGAGGTGTCCAAAATTGCCAACATCTACCTGGAAGA 335  
DB 248 ACCTGCATCATGTGGGGAAGAGGTGTCGAAATCGGGAAGCATTTGCCGGGAGGA 307  
QY 336 CTGATATGATCAAGACATATTGAGAGGACAGATCCAGAACATCA 386  
DB 308 CCGACAAATGAATTAAGAACTACTGAGAGGACCAAGATCCAAAGCAGATTA 358

RESULT 5

ID AAC41714 standard; DNA; 621 BP.

XX AAC41714;

DT 17-OCT-2000 (first entry)

DE Arabidopsis thaliana DNA fragment SEQ ID NO: 32868.

KW Hybridisation assay; genetic mapping; gene expression control;

KW protein identification; signal transduction pathway;

OS Arabidopsis thaliana.

PN EP1033405-A2.

PD 06-SHP-2000.

PF 25-FEB-2000; 2000EP-0301439.

PR 25-FEB-1999; 99US-0121825.

PR 05-MAR-1999; 99US-0123180.

PR 09-MAR-1999; 99US-0123548.

PR 23-MAR-1999; 99US-0125788.

PR 25-MAR-1999; 99US-0126264.

PR 29-MAR-1999; 99US-0126785.  
PR 01-APR-1999; 99US-0127462.  
PR 06-APR-1999; 99US-0128234.  
PR 08-APR-1999; 99US-0128714.  
PR 16-APR-1999; 99US-0129845.  
PR 19-APR-1999; 99US-0130077.  
PR 21-APR-1999; 99US-0130449.  
PR 23-APR-1999; 99US-0130510.  
PR 23-APR-1999; 99US-0130891.  
PR 28-APR-1999; 99US-0131449.  
PR 30-APR-1999; 99US-0132048.  
PR 30-APR-1999; 99US-0132407.  
PR 04-MAY-1999; 99US-0132487.  
PR 05-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.  
PR 06-MAY-1999; 99US-0132487.  
PR 07-MAY-1999; 99US-0132863.  
PR 11-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 14-JUN-1999; 99US-0138847.  
PR 16-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.

PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144332.  
PR 20-JUL-1999; 99US-0144332.  
PR 20-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 28-JUL-1999; 99US-0145919.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 05-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147302.  
PR 06-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147316.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151348.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158232.

PR 12-OCT-1999; 99US-0158369.  
PR 13-OCT-1999; 99US-0159293.  
PR 13-OCT-1999; 99US-0159294.  
PR 13-OCT-1999; 99US-0159295.  
PR 14-OCT-1999; 99US-0159329.  
PR 14-OCT-1999; 99US-0159330.  
PR 14-OCT-1999; 99US-0159331.  
PR 14-OCT-1999; 99US-0159637.  
PR 18-OCT-1999; 99US-0159638.  
PR 21-OCT-1999; 99US-0160741.  
PR 21-OCT-1999; 99US-0160767.  
PR 21-OCT-1999; 99US-0160770.  
PR 21-OCT-1999; 99US-0160814.  
PR 21-OCT-1999; 99US-0160815.  
PR 22-OCT-1999; 99US-0160980.  
PR 22-OCT-1999; 99US-0160981.  
PR 22-OCT-1999; 99US-0160989.  
PR 25-OCT-1999; 99US-0161404.  
PR 25-OCT-1999; 99US-0161405.  
PR 26-OCT-1999; 99US-0161359.  
PR 26-OCT-1999; 99US-0161360.  
PR 26-OCT-1999; 99US-0161361.  
PR 28-OCT-1999; 99US-0161920.  
PR 28-OCT-1999; 99US-0161922.  
PR 28-OCT-1999; 99US-0161993.  
PR 29-OCT-1999; 99US-0162142.

Query Match 24.1%; Score 208; DB 21; Length 621;  
Best Local Similarity 77.1%; Pred. No. 1,8e-32;  
Matches 253; Conservative 0; Mismatches 75; Indels 0; Gaps 0;

QY 60 AAGATCCTGAATGAGAAAGGCGCTTGACATGGAAGAGACTTGATCTTGATGAAC 119  
DB 62 AAGAGGAACAGTGAAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 121  
QY 120 AATATGCAATCATGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 179  
DB 122 ACATCCTTATCATGAG 181  
QY 180 GTACGGAAG 239  
DB 182 GTACTGGAAG 241  
QY 240 GGAATTTTACACCCGAG 299  
DB 242 GGAACATTAACCGAG 301  
QY 300 GGTGTCGCAAAATGCGACAGATCTACTGGAAGAGAGAGAGAGAGAGAGAGAGAGAG 359  
DB 302 GGTGTCGCAAAATGCGAG 361  
QY 360 GGAGGACAAAGATCCAG 387  
DB 362 GGAGGACAAAGATTCAG 389

RESULT 6  
AAC37953  
ID AAC37953 standard; DNA; 1127 BP.  
XX AAC37953;  
AC  
XX  
DT 17-OCT-2000 (first entry)  
XX  
XX Arabidopsis thaliana DNA fragment SEQ ID NO: 19258.  
DE  
XX  
XX Hybridisation assay; genetic mapping; gene expression control;  
KW protein identification; signal transduction pathway;  
KW metabolic pathway; promoter; termination sequence; ss.  
XX

[illegible]



PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158232.  
PR 12-OCT-1999; 99US-0158369.  
PR 13-OCT-1999; 99US-0159293.  
PR 13-OCT-1999; 99US-0159294.  
PR 13-OCT-1999; 99US-0159295.  
PR 14-OCT-1999; 99US-0159329.  
PR 14-OCT-1999; 99US-0159330.  
PR 14-OCT-1999; 99US-0159331.  
PR 14-OCT-1999; 99US-0159637.  
PR 18-OCT-1999; 99US-0159638.  
PR 18-OCT-1999; 99US-0159584.  
PR 21-OCT-1999; 99US-0160741.  
PR 21-OCT-1999; 99US-0160767.  
PR 21-OCT-1999; 99US-0160768.  
PR 21-OCT-1999; 99US-0160770.  
PR 21-OCT-1999; 99US-0160814.  
PR 21-OCT-1999; 99US-0160815.  
PR 22-OCT-1999; 99US-0160980.  
PR 22-OCT-1999; 99US-0160981.  
PR 22-OCT-1999; 99US-0160989.  
PR 25-OCT-1999; 99US-0161404.  
PR 25-OCT-1999; 99US-0161405.  
PR 25-OCT-1999; 99US-0161406.  
PR 26-OCT-1999; 99US-0161359.  
PR 26-OCT-1999; 99US-0161360.  
PR 26-OCT-1999; 99US-0161361.  
PR 28-OCT-1999; 99US-0161920.  
PR 28-OCT-1999; 99US-0161992.  
PR 28-OCT-1999; 99US-0161993.  
PR 29-OCT-1999; 99US-0162142.

Query Match 24.1%; Score 208; DB 21; Length 1127;  
Best Local Similarity 77.1%; Pred. No. 1.9e-32;  
Matches 253; Conservative 0; Mismatches 75; Indels 0; Gaps 0;

QY 60 AAGATCCTGAAGTGAAGAAAGGCGCTTGACATGGAAGAAGACTTGATCTTGATGAACT 119  
DB 335 AAGAAGGACGATGAGAAAGACCTTGACATGGAAGAAGATTCATCCTCTTTAATT 394  
QY 120 ATATTGCAATCATGGGAAGGTGTTGGAACCTTTGGCCAAAGCTGCTGCTCAAACT 179  
DB 395 ACATCTTAATCATGATGGAAGGTGTTGGAACCTGCTGCTGCAAAAGCCTGCTGCTTAAC 454  
QY 180 GTAACGGAAGAGTGGCCGGCTAAGTGCTAAATTAACCTCGCTCGATGTTAGAAAG 239  
DB 455 GTACTGCAAAAAGTTGCGGCTCGGCTGCTGGAAGTATCTCCGACCAAGTGTGCGGCGAG 514  
QY 240 GGAATATTACACCGGAGGAACAATTTTGATTATGAGAGCTCCAGCAAAAGGGGGAACA 299  
DB 515 GGAACATTAACACAGAAAGAAAGCTTTTGATCATTCAGCTTCATGCTTAAGCTTGGAACA 574  
QY 300 GGTGTCGCAAAATTTGCCAAGCATCTACTGGAAGAGACTGATTAATGAGATCAAGAACTATT 359  
DB 575 GGTGTCGCAAAATTTGCCAAGCATCTACTGGAAGAGACTGATTAATGAGATCAAGAACTATT 634  
QY 360 GGAGACAGAGATCCAGACATCA 387  
DB 635 GGAGACAGAGATTCAGACATCA 662

RESULT 7

AAH87724  
ID AAH87724 standard; cDNA; 626 BP.  
XX  
AC AAH87724;  
XX  
DT 25-SEP-2001 (first entry)  
XX  
DE Peppermint plant oil gland expressed cDNA 80.  
XX  
KW Peppermint; plant oil gland cell; terpenoid essential oil; resin;  
XX genetic mapping; antisense suppression; recombinant expression; ss.  
OS Mentha x piperita.  
XX  
PN WO200153319-A1.  
XX  
PD 26-JUL-2001.  
XX  
PF 19-JAN-2001; 2001WO-US02567.  
XX  
PR 20-JAN-2000; 2000US-0177264.  
XX  
PA (CROT/) CROTEAU R. B.  
PA (LANG/) LANGE B. M. R.  
PA (WILD/) WILDUNG M. R.  
XX  
PI Croteau RB, Lange BM, Wildung MR;  
XX  
DR WPI; 2001-488706/53.  
XX  
PT New nucleic acid molecules corresponding to mRNA molecules expressed in  
XX peppermint oil glands for enhancing expression of plant oil gland cell  
XX proteins -  
XX  
PS Claim 1; Page 108; 251pp; English.  
XX  
CC The invention relates to nucleic acid molecules (AAH87645-AAH88116) that  
XX correspond to all or part of a mRNA molecule expressed in plant oil  
XX gland cells, especially peppermint and plant oil glands that produce  
XX terpenoid essential oils and resins. The nucleic acids are useful for  
XX genetically mapping a plant genome for genes expressed in plant oil  
XX gland cells and to suppress (for example by antisense suppression) or  
XX enhance their expression (for example by genetically transforming a  
XX plant cell with a replicable expression vector that expresses one or more  
XX CC proteins naturally expressed in plant oil gland cells). The nucleic acids  
XX are also useful for recombinant expression of plant oil gland proteins  
XX required for terpenoid essential oil and/or resin production in bacterial  
XX and/or yeast cells.  
CC  
SQ Sequence 626 BP; 187 A; 158 C; 163 G; 118 T; 0 other;  
XX  
Query Match 20.3%; Score 175; DB 22; Length 626;  
Best Local Similarity 72.7%; Pred. No. 6.6e-26;  
Matches 226; Conservative 0; Mismatches 85; Indels 0; Gaps 0;  
QY 68 GAAGTGAAGAAAGGCGCTTGACATGGAAGAAGACTTGATCTTGATGAACTATGCA 127  
DB 155 GAGGTGGAAGAGGCGCTTGACATGGAAGAAGACTTGATCTTGATGAACTATGCA 214  
QY 128 AATCATGGGGAAGGTGTTGGAACCTTTGGCCAAAGCTGCTGCTCAAAAGTAAACGA 187  
DB 215 CATCAGGCGCAAGGCCGATGGAAGCTCTGCGCTCGCTCAGCAAGGCTCAACAGAACTGGA 274  
QY 188 AAGAGTTCGCGGCTAAGTGCTAAATTAACCTCGCTGATGTTAGAAAGAGGAATATT 247  
DB 275 AAGAGTTCGCGGCTAAGTGCTAAATTAACCTCGCTGATGTTAGAAAGAGGAATATT 334  
QY 248 ACACCCGAGGAACAATTTTGATGAGAGCTCCAGCAAAAGTGGGGAACAGGTGTC 307  
DB 335 ACTCTGGAAGAGCAAGCTTTGATTCGACCTTCGATGAGGTGGGCAACAGGTGTC 394  
QY 308 AAAATGCCAAGCATCTACTGGAAGAGACTGATTAATGAGATCAAGAACTATTGAGGACA 367

Db 395 AAAATCGCGCAGCATCTGCCGGAGAACTGACTACGAATAAAGACTACTGAGAACCA 454  
 QY 368 AGATCCAGAA 378  
 || || ||  
 Db 455 AGAGTGCAAAA 465

RESULT 8  
 AAD05784  
 ID AAD05784 standard; cDNA: 1137 BP.  
 XX  
 AC AAD05784;  
 XX  
 DF 31-JUL-2001 (first entry)  
 XX  
 DE Arabidopsis thaliana transcription factor, G1324 cDNA.  
 XX  
 KW Transcription factor; biochemical characteristic; controlling element;  
 KW structural characteristic; developmental characteristic; gene therapy;  
 KW agricultural biotechnology; plant trait modification; ss.  
 XX  
 OS Arabidopsis thaliana.  
 XX  
 FT Key Location/Qualifiers  
 FT CDS 54..914  
 FT /\*tag= a  
 FT /product= "Transcription factor, G1324"

MO200136597-A1.  
 PD 25-MAY-2001.  
 XX  
 PE 14-NOV-2000; 2000WO-US31344.  
 XX  
 PR 17-NOV-1999; 990US-0166228.  
 PR 17-A2R-2000; 2000US-0197899.  
 PR 22-AUG-2000; 2000US-0227439.  
 XX  
 PA (MENDEL-) MENDEL BIOTECHNOLOGY INC.  
 PA (YUG3/) YU G.  
 PA (ADAM/) ADAM L.  
 PA (RIB2/) RIECHMANN J L.  
 PA (HBA3/) HEARD J.  
 PA (SAMA/) SAMAH R.  
 PA (PI3/) PILGRIM M.  
 PA (PINE/) PINE O.  
 PA (JIANG/) JIANG C.

PI Creelman R, Yu G, Adam L, Riechmann JL, Heard J, Samaha R;  
 PI Pilgrim M, Pineda O, Jiang C;  
 DR WPI: 2001-335999/35.  
 DR P-PSDB; AAE01900.  
 XX  
 XX  
 PT Nucleic acids encoding plant transcription factor polypeptides, useful  
 PT for altering the biochemical characteristics of plants e.g. corn,  
 PT potato and cotton plants -  
 XX  
 XX  
 PS Claim 4; Page 87-88; 127PP; English.

The present sequence is Arabidopsis thaliana transcription factor, G1324 cDNA. The transcription factor is used for altering a plant's biochemical characteristics. The transcription factor may be used to alter the structure and developmental characteristics of plants such as soybean, wheat, corn, potato, cotton, rice, oilseed rape, sunflower, alfalfa, sugar cane, turf, banana, blackberry, blueberry, strawberry, raspberry, cantaloupe, carrot, cauliflower, coffee, cucumber, eggplant, grapes, honey dew, lettuce, mango, melon, onion, papaya, peas, peppers, pineapple, spinach, squash, sweet corn, tobacco, tomato, watermelon, rosaceous fruits and/or vegetable brassicas. Transcription factors are key controlling elements of biological pathways and altering expression levels of 1 or more transcription factors can change entire biological

CC pathways in an organism. Therefore manipulating transcription factor  
 CC levels in plants offers great potential in agricultural biotechnology  
 CC for modifying a plant's traits. Transcription factor cDNA is useful in  
 CC gene therapy.  
 XX  
 SQ Sequence 1137 BP; 378 A; 213 C; 233 G; 313 T; 0 other;

Query Match 19.8%; Score 170.6; DB 22; Length 1137;  
 Best Local Similarity 67.0%; Pred. No. 5.3e-25;  
 Matches 242; Conservative 0; Mismatches 119; Indels 0; Gaps 0;

QY 21 ACAAGTCATGATGATAAAAAACAACAGTGTAAAGCTGCTCAAGTCTGAAGTGAAGAAAG 80  
 Db 58 AAAATTCGATGAAGAAAGAAAGAAAGAGCTTCAAGAAAGTGAAGTGAAGTGAAGAAAG 117  
 QY 81 GGCCCTTGACATGAAGAAAGAAAGTGTATGATGATGATGATGATGATGATGATGATGATG 140  
 Db 118 GGCCCTTGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 177  
 QY 141 GTGTTTGGAACCTTTGGCCAAAGCTGCTGCTCAACGTAACGGAAGAGTTGCCGCGC 200  
 Db 178 GTGTTTGGAATCACGTCGCCAAATGTGCGGCTTAAAGAACTGGGAAAGTTGTAGAT 237  
 QY 201 TAAGTGCGCTAAATTAATCTCCCTGATGATGATGATGATGATGATGATGATGATGATG 260  
 Db 238 TGAGATGCTTGAATTAATCTCCCTGATGATGATGATGATGATGATGATGATGATGATG 297  
 QY 261 AACTTTGATTAATGAGCTCCACGCAAAAGTGGGGAAGAGTGTCTCAAAATTTGCCAGC 320  
 Db 298 AGCTTTGATCTCTGAGCTTCAATTAATGAGTGAATGAGTGAATGAGTGAATGAGTGAAT 357  
 QY 321 ATCTACCTGGAAGACTGATTAATGAGTCAAGAACTATTGGAGACAGATCCAGAAAGC 380  
 Db 358 ACTTGCCAGGAAGAAAGAGATTAAGAGATCAAGAACTATTGGAGAAAGAGTTCAAAAAC 417  
 QY 381 A 381  
 Db 418 A 418

RESULT 9  
 AAX25572  
 ID AAX25572 standard; cDNA: 837 BP.  
 XX  
 AC AAX25572;  
 XX  
 DT 02-AUG-1999 (first entry)  
 XX  
 DE Arabidopsis thaliana MYB2 (AtMYB2) cDNA.  
 XX  
 KW MYB2 gene; AtMYB2 gene; transcription factor; transgenic plant;  
 KW environmental stress; stress tolerance; anaerobic stress;  
 KW drought; flooding; salt; cold; crop protection; ss.  
 XX  
 OS Arabidopsis thaliana.  
 XX  
 FT Key Location/Qualifiers  
 FT CDS 10..831  
 FT /\*tag= a  
 XX  
 XX  
 PN WO9916878-A1.  
 XX  
 PD 08-APR-1999.  
 XX  
 PF 25-SEP-1998; 98WO-AU00812.  
 XX  
 PR 26-SEP-1997; 97AU-0009479.  
 XX  
 PA (CSIR ) COMMONWEALTH SCI & IND RES ORG.  
 PI Dennis ES, Dolferus RM, Hoeren FU, Peacock WJ.  
 DR WPI: 1999-263695/22.

DR P-PSDB: AAY05831.  
XX Altering expression of a stress-related gene in a plant  
XX Claim 22; Page 72-74; 119pp; English.  
XX This is the sequence of Arabidopsis thaliana ecotype C24 cDNA  
XX encoding the stress-related protein AtMYB2 (see AAY05831). The  
XX cDNA was obtained by RT-PCR of anaerobically induced root RNA  
XX derived from ecotype C24. The inventors have demonstrated that a  
XX family of transcription factors, the MYB2 family of proteins, are  
XX capable of inducing the expression of stress-related genes in plant  
XX cells in response to stress situations. The MYB2 protein encoded  
XX by the AtMYB2 gene is useful for the purposes of inducing or  
XX repressing the expression of plant-expressible genes that are  
XX involved in the plant's response to anaerobic stress, flooding  
XX stress, cold stress, dehydration stress, drought stress, heat  
XX stress or salinity. The stress-related gene, such as the MYB2  
XX alcohol dehydrogenase Adh1 gene, is transactivated by the MYB2  
XX cis-acting regulatory sequence, in particular a MB5-1 (Myb binding  
XX site) motif and/or a GC-motif and/or a G-box-1 motif and/or a  
XX G-box-2 motif, as found in the Adh1 promoter (see AAX25573-74). The  
XX MYB2 protein is able to target several stress-induced enzymes,  
XX rather than targeting just one, allowing the production of stress  
XX tolerant transgenic plants.  
SQ Sequence 837 BP; 246 A; 176 C; 212 G; 203 T; 0 other;  
Query Match 19.6%; Score 169.4; DB 20; Length 837;  
Best Local Similarity 69.5%; Pred. No. 8; 8e-25;  
Matches 230; Conservative 0; Mismatches 101; Indels 0; Gaps 0;  
QY 60 AAGATCTGAGAGGAAAGGCGCTTGACAAATGGAAGAAGACTGATGTAAGT 119  
DB 56 AAGATTCTGATGACGGAAGAGTCCATGACCGAAGAAGATGCAATCTAGTCAACT 115  
QY 120 ATATTGCAATTCATGGGAGGAGTGTGTAACCTTTGGCCAAAGCTGCTGCTCAAC 179  
DB 116 TCGCTCTATTTCATGCGGATGCTGCTGGAACACATCGCTGCTTCCCTGGGCTAAAGC 175  
QY 180 GTAACGGAAAGAGTGGCGGCTAAGTGGCTAAATTAATCACTCGCTCGATGTAAGAAG 239  
DB 176 GAAGTGGTAGAGTGTGATTAAGATGGCTTAATTAATCACTCGATGTAAGAAG 235  
QY 240 GGAATATTACACCCGAGGAACAATTGATTAATGAGACTCCACGCAAAAGTGGGAACA 299  
DB 236 GCAACATCACTCTCGAAGAAACAATTAATGCTCAAACTCTTTGGGCAATTA 295  
QY 300 GGTGCTCAAAATTTGCCAAGCATCTACCTGGAAGAGCATGATTAATGATCAAACTATT 359  
DB 296 GTGCTGGAAGATTGGCCATATCTACCGGAAGAACAGATTAATGAATTAATTA 355  
QY 360 GGAAGACAAAGATCCAGAACACATCAAGA 390  
DB 356 GGAGAACTCGATCAAAAGCAAGCAACA 386  
RESULT 10  
ABK65252  
ID ABK65252 standard: cDNA; 981 BP.  
XX  
AC ABK65252;  
XX  
DT 02-JUL-2002 (first entry)  
XX  
DE Arabidopsis cDNA encoding a transcription factor #104.  
XX  
XX Plant; ss; gene; transcription factor; transgenic;  
KW agriculture; metabolic chemical; environmental stress; drought;  
KW microbial disease resistance; herbicide resistance; seed yield;  
KW fruit yield; growth rate; leaf senescence; flower senescence.  
XX

OS Arabidopsis thaliana.  
XX  
XX WO200215675-A1.  
XX  
XX 28-FEB-2002.  
XX  
XX 22-AUG-2001; 2001WO-US26189.  
XX  
XX 22-AUG-2000; 2000US-227439P.  
XX  
XX 16-NOV-2000; 2000US-0713994.  
XX  
XX 16-APR-2001; 2001US-0837944.  
XX  
XX (MEND-) MENDEL BIOTECHNOLOGY INC.  
XX  
XX (PIG/) PILGRIM M.  
XX  
XX (CREE/) CREELMAN R.  
XX  
XX (DUBE/) DUBBEL A J.  
XX  
XX (HEAR/) HEARD J.  
XX  
XX (JIAN/) JIANG C.  
XX  
XX (KEDD/) KEDDIE J.  
XX  
XX (ADAM/) ADAM L.  
XX  
XX (RATC/) RATCLIFF O.  
XX  
XX (REUB/) REUBER J L.  
XX  
XX (RIEC/) RIECHMANN J L.  
XX  
XX (YUGG/) YU G.  
XX  
XX (PINE/) PINEDA O.  
XX  
XX Pilgrim M, Creelman R, Dubell AJ, Jiang C, Keddie J,  
XX Adam L, Ratcliff O, Reuber JL, Riechmann JL, Yu G, Pineda O;  
XX WPI: 2002-292022/33.  
XX P-PSDB: AA093066.  
XX  
XX An isolated or recombinant polynucleotide used to produce a transgenic  
XX plant -  
XX  
XX Claim 4; Page 443-445; 941pp; English.  
XX  
XX The invention relates to 1 of 232 isolated or recombinant polynucleotides  
XX encoding an Arabidopsis thaliana transcription factor, their variants,  
XX complements, fragments, or related polynucleotide with 31% to 95%  
XX sequence identity, where the plant possesses an altered trait as compared  
XX to a wild-type or reference plant, or the plant exhibits an altered  
XX phenotype as compared to a wild-type or reference plant, or the plant  
XX exhibits ectopic expression or altered expression of one or more genes  
XX associated with a plant trait as compared to a wild plant. Also included  
XX are a transgenic plant comprising the polynucleotides, a computer  
XX readable medium having stored sequence information, and identifying a  
XX homologue sequence from a database comprising a plurality of known plant  
XX sequences comprising inputting sequence information selected from one of  
XX 464 fully defined sequences given in the specification. The isolated or  
XX recombinant polynucleotide is used for producing a plant having a  
XX modified trait, the method comprising selecting a polynucleotide that  
XX encodes a polypeptide or an antisense nucleic acid, inserting the  
XX polynucleotide or antisense nucleic acid into an expression vector,  
XX introducing the vector into a plant or a cell of a plant to overexpress  
XX the polypeptide or antisense nucleic acid, thereby producing a modified  
XX plant, and selecting for a modified trait (e.g. increased  
XX production of agriculturally useful proteins or metabolic chemicals,  
XX pest tolerance, environmental stress response (e.g. drought), microbial  
XX disease resistance, herbicide resistance, seed and fruit yield, growth  
XX rate, leaf and flower senescence and many other traits listed in the  
XX specification). The present sequence is one of the 232 polynucleotides  
XX encoding an A. thaliana transcription factor.  
XX  
SQ Sequence 981 BP; 319 A; 193 C; 235 G; 234 T; 0 other;  
Query Match 19.6%; Score 169.4; DB 24; Length 981;  
Best Local Similarity 69.5%; Pred. No. 9e-25;  
Matches 230; Conservative 0; Mismatches 101; Indels 0; Gaps 0;  
QY 60 AAGATCTGAGAGGAAAGGCGCTTGACAAATGGAAGAAGACTGATGTAAGT 119  
DB 56 AAGATTCTGATGACGGAAGAGTCCATGACCGAAGAAGATGCAATCTAGTCAACT 159



PA (PINE/) PINEDA O.  
XX  
PI Pilgrim M, Creelman R, Dubeil AJ, Heard J, Jiang C, Keddle J;  
PI Adam L, Ratcliff O, Reuber JL, Riechmann JL, Yu G, Pineda O;  
XX  
DR WPI: 2002-292022/33.  
DR P-PSDB: AAU93178.  
XX  
PT An isolated or recombinant polynucleotide used to produce a transgenic  
PT plant -  
XX  
PS Claim 4: Page 881-883; 941pp; English.  
XX  
CC The invention relates to 1 of 232 isolated or recombinant polynucleotides  
CC encoding an Arabidopsis thaliana transcription factor, their variants,  
CC complements, fragments, or related polynucleotide with 31% to 95%  
CC sequence identity, where the plant possesses an altered trait as compared  
CC to a wild-type or reference plant, or the plant exhibits an altered  
CC phenotype as compared to a wild-type or reference plant, or the plant  
CC exhibits ectopic expression or altered expression of one or more genes  
CC associated with a plant trait as compared to a wild plant. Also included  
CC are a transgenic plant comprising the polynucleotides, a computer  
CC readable medium having stored sequence information, and identifying a  
CC homologue sequence from a database comprising a plurality of known plant  
CC sequences comprising inputting sequence information selected from one of  
CC 464 fully defined sequences given in the specification. The isolated or  
CC recombinant polynucleotide is used for producing a plant having a  
CC modified trait, the method comprising selecting a polynucleotide that  
CC encodes a polypeptide or an antisense nucleic acid, inserting the  
CC polynucleotide or antisense nucleic acid into an expression vector,  
CC introducing the vector into a plant or a cell of a plant to overexpress  
CC the polypeptide or antisense nucleic acid, thereby producing a modified  
CC plant, and selecting for a modified trait (e.g. increased  
CC production of agriculturally useful proteins or metabolic chemicals,  
CC pest tolerance, environmental stress response (e.g. drought), microbial  
CC disease resistance, herbicide resistance, seed and fruit yield, growth  
CC rate, leaf and flower senescence and many other traits listed in the  
CC specification). The present sequence is one of the 232 polynucleotides  
CC encoding an A. thaliana transcription factor.  
XX  
SQ Sequence 972 BP; 318 A; 217 C; 216 G; 221 T; 0 other;  
XX  
Query Match 18.9%; Score 163.4; DB 24; Length 972;  
Best Local Similarity 67.4%; Pred. No. 1.4e-23;  
Matches 230; Conservative 0; Mismatches 111; Indels 0; Gaps 0;  
XX  
QY 50 AAGAGCTCTCAAGATCCTGAAGTGAAGAAAGCGCTTGACCAATGGAAGAACTTATC 109  
DB 34 AACATGGAAGACGAGTGAAGTGAAGAGAGTCCGTGAGCTGTGAAGAAAGATTTTAAG 93  
QY 110 TTGATGAACATTAATGCAAAATCATGGGAAGGTGTTGAACTCTTTGGCCAAAGCTGCT 169  
DB 94 CTCATGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 153  
QY 170 GGTCCTCAACCTAAGCGAAGAGTGGCGGCTAAGTGGCTAATTAATTAATTAATTAAT 229  
DB 154 GGCTTCCAAACCCACCGGTAAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 213  
QY 230 GTTGAAGAGGGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 289  
DB 214 GTCCGCGGTGAAACATTACCTGTTGAAGAACATCTTGATCTCGAATCTTATCCCGT 273  
QY 290 TGGGGAAGAGAGTGTCCAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 349  
DB 274 TGGGGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 333  
QY 350 AAGAATCTTTGGAGCAAGATCCAGAACATCAACGA 390  
DB 334 AAGAATCTTTGGAGCAAGATCCAGAACATCAACGA 374  
XX  
RESULT 13  
AAC56152

ID AAC56152 standard; DNA; 389 BP.  
XX  
AC AAC56152;  
XX  
DT 25-JAN-2001 (first entry)  
XX  
DE Eucalyptus grandis transcription factor DNA sequence #283.  
XX  
XX Plant; transcription factor; gene expression; eucalyptus; pine; acacia;  
KW poplar; sweetgum; teak; mahogany; bZIP; G-box binding factor;  
KW basic helix-loop-helix zipper; homeotic; homeodomain; homeobox; MADS;  
KW homeodomain zipper; LIM domain; AP2; ERES; zinc finger domain;  
KW type 2 Cys2His2; CCAAT box element; MIB; ss.  
XX  
OS Eucalyptus grandis.  
XX  
PN WO200053724-A2.  
XX  
PD 14-SEP-2000.  
XX  
PF 09-MAR-2000; 2000WO-US06112.  
XX  
PR 11-MAR-1999; 99US-0266513.  
PR 18-AUG-1999; 99US-0149485.  
XX  
PA (GENE-) GENESIS RES & DEV CORP LTD.  
PA (FLET-) FLETCHER CHALLENGE FORESTS LTD.  
PI Wood M, McGrath A, Shenk MA, Glenn M;  
XX  
DR WPI: 2000-579369/54.  
XX  
PT New isolated polynucleotide encoding a plant transcription factor for  
PT producing a plant e.g. a woody plant, preferably eucalyptus or pine,  
PT having modified gene expression or modified activity of a polypeptide  
PT -  
XX  
PS Claim 1; Page 120; 747pp; English.  
XX  
XX  
CC The present invention relates to novel plant transcription factors from  
CC Eucalyptus grandis or Pinus radiata. The present sequence is the coding  
CC sequence for one such transcription factor. The transcription factor may  
CC be used to produce a plant having modified gene expression such as a  
CC woody plant e.g. a eucalyptus, pine, acacia, poplar, sweetgum, teak, or  
CC mahogany species or to modify the activity of a polypeptide in a plant.  
CC The transcription factors of the present invention are members from the  
CC following families of regulatory proteins: bZIP, bZIP family of G-box  
CC binding factors, basic helix-loop-helix zipper, LIM domain, AP2  
CC homeotic/homeodomain/homeobox/MADS, homeodomain zipper, LIM domain, AP2  
CC and ERES, zinc finger domains of type 2 Cys2His2, CCAAT box elements  
CC and MYB.  
XX  
SQ Sequence 389 BP; 112 A; 102 C; 97 G; 78 T; 0 other;  
XX  
Query Match 18.7%; Score 161.6; DB 21; Length 389;  
Best Local Similarity 71.6%; Pred. No. 2.9e-23;  
Matches 212; Conservative 0; Mismatches 84; Indels 0; Gaps 0;  
XX  
QY 8 CTCATCAACACACAAAGTCAATGATTAATAAACAACAGTGAAGCTGTCAAGATCCT 67  
DB 29 CTCTCCAAAGTGAACATGACAAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGTGTG 88  
QY 68 GAAGTGAAGAAAGGCGCTTGGAACATGAAGAGACTTGTGATGATGATGATGATGATGAT 127  
DB 89 GAAGTGAAGAAAGGCGCTTGGAACATGAAGAGACTTGTGATGATGATGATGATGATGAT 148  
QY 128 AATCATGAGGGAAGGTGTTGGAACCTTTGGCCAAAGTGTGCTCAACAGTAACGA 187  
DB 149 AATCATGAGGGAAGGCGCTTGGAACCTTTGGCCAAAGTGTGCTCAACAGTAACGA 208  
QY 188 AAGAGTTGGCGGCTGAAGTGAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 247  
DB 209 AAGAGTTGGCGGCTGAAGTGAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 268





1  
2  
3  
4

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65  
66  
67  
68  
69  
70  
71  
72  
73  
74  
75  
76  
77  
78  
79  
80  
81  
82  
83  
84  
85  
86  
87  
88  
89  
90  
91  
92  
93  
94  
95  
96  
97  
98  
99  
100



GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 18, 2003, 07:57:25 ; Search time 1682 Seconds

(without alignments)  
8309.575 Million cell updates/sec

Title: US-10-021-811-35

Perfect score: 863  
Sequence: 1 gcaagagctctatcacacac.....aaaaaaaaaaaaaaaaa 863

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :

EST:\*  
1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estin:\*  
4: em\_estinu:\*  
5: em\_estov:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_hic:\*  
9: gb\_estl:\*  
10: gb\_estl2:\*  
11: gb\_hic:\*  
12: gb\_est3:\*  
13: gb\_est4:\*  
14: gb\_est5:\*  
15: em\_estfun:\*  
16: em\_estlom:\*  
17: gb\_gsa:\*  
18: em\_gsa\_hum:\*  
19: em\_gsa\_hiv:\*  
20: em\_gsa\_pln:\*  
21: em\_gsa\_vrt:\*  
22: em\_gsa\_fun:\*  
23: em\_gsa\_mam:\*  
24: em\_gsa\_mus:\*  
25: em\_gsa\_other:\*  
26: em\_gsa\_pro:\*  
27: em\_gsa\_tod:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	562	65.1	562	14	BM732121 sa173b11.
2	506.4	58.7	586	12	BE820766 GM700012A
3	491.4	56.9	782	10	BE658316 BM527606
4	462.6	53.6	501	13	BM527606 sa163906
5	444.4	51.5	450	10	BE058947 sn23b01.Y
6	442.8	51.3	523	9	AI930997 sb45h07.Y

7	425.4	49.3	431	10	AM459279
8	424.4	49.2	552	13	BM527774
9	401.2	46.5	568	13	BM527664
10	400.6	46.4	544	13	BM527508
11	369.8	42.9	409	14	BE804790
12	369.6	42.8	560	14	BM732539
13	368.6	42.7	635	14	BO146360
14	363.6	42.1	476	10	BE057370
15	347	40.2	542	13	BM528383
16	330.8	38.3	425	10	AM423958
17	307.8	35.7	682	14	BO146831
18	306.2	35.5	688	13	BI272897
19	304.8	35.3	669	12	BC457971
20	304.6	35.3	395	12	BE805071
21	294	34.1	470	10	AM156542
22	291.2	33.7	447	13	BI273011
23	288.4	33.4	558	10	BE324639
24	284	32.9	766	10	BE659054
25	279.4	32.4	715	10	BE658737
26	279	32.3	567	14	BM732162
27	275	31.9	592	10	AM928296
28	271.6	31.5	500	9	AI897784
29	271.6	31.5	519	9	AI897681
30	265.6	30.8	397	10	AM432364
31	265.4	30.8	681	9	AI486576
32	258.8	30.0	503	14	BO104458
33	246	28.5	502	14	BO106505
34	240.2	27.8	463	9	AI938482
35	223.8	25.9	516	12	BE652172
36	219.2	25.4	467	12	BE807621
37	217.8	25.2	370	9	AI487424
38	217.8	25.2	371	9	AI487923
39	204.8	23.7	542	9	AU238946
40	193	22.4	492	13	BI262794
41	192.4	22.3	480	10	BE023108
42	190.2	22.0	343	12	BE804882
43	189.4	21.9	256	12	BE805198
44	184.8	21.4	510	12	BF325282
45	183.6	21.3	295	13	BM527578

ALIGNMENTS

RESULT 1  
BM732121  
LOCUS 562 bp mRNA linear EST 01-MAR-2002  
DEFINITION sa173b11.Y1 Gm-c1061 glycine max cDNA clone SOYBEAN CLONE ID:  
Gm-c1061-4678 5' similar to TR:Q39028 Q39028 ATMYB2. ; mRNA  
sequence.  
ACCESSION BM732121  
VERSION BM732121.1 GI:19053454  
KEYWORDS EST.  
SOURCE soybean.  
ORGANISM Glycine max  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;  
Glycine.

REFERENCE 1 (bases 1 to 562)  
AUTHORS Shoemaker,R., Keim,P., Vodkin,L., Expeiding,J., Corryell,V., Khana  
A., Bolla,B., Marita,M., Hillier,L., Kucaba,T., Martin,J., Beck,C.,  
Wyllie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers  
Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk  
R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann  
R., Matternson,R. and Wilson,R.

TITLE Public Soybean EST Project  
JOURNAL Unpublished (1999)  
COMMENT Contact: Shoemaker R/Public soybean EST Project  
Public Soybean EST Project  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA  
Tel: 314 286 1800

Fax: 314 286 1810  
 Email: est@watson.wustl.edu  
 This clone is available through: ResGen, Invitrogen Corp. 2130  
 South Memorial Parkway Huntville, AL 35801 For further information  
 call: (800)-533-4363 or contact: cou@resgen.com web site:  
 www.resgen.com  
 Seq primer: -40RP from Glibco  
 High quality sequence stop: 421.  
 Location/Qualifiers

# FEATURES

source

1. 562  
 /organism="Glycine max"  
 /db\_xref="taxon:3847"  
 /clone="SOYBEAN CLONE ID: Gm-cl061-4678"  
 /clone\_1lb="Gm-cl061"  
 /tissue\_type="mature flowers of field grown plants"  
 /lab\_host="DH10B"  
 /note="Vector: Bluescript II SK<sup>+</sup>; Site\_1: EcoRI; Site\_2:  
 XhoI; The cDNA library was constructed from mRNA isolated  
 from mature flowers of field grown plants for the cultivar  
 Raiden. Complementary DNA was synthesized from mRNA using  
 a primer consisting of a poly(dT) sequence with a XhoI  
 restriction site. EcoRI adapters were ligated to the  
 blunt-ended cDNA fragments followed by XhoI digestion. The  
 cDNA fragments were directionally cloned into the  
 EcoRI-XhoI restriction site of the Bluescript vector. The  
 ligated cDNA fragments were transformed into DH10B host  
 cells (GibcoBRL). This library was constructed in the  
 laboratory of Dr. Randy Shoemaker."

BASE COUNT  
 ORIGIN

187 a 115 c 128 g 132 t

## Query Match

Best Local Similarity 100.0%; Pred. NO. 2.8e-74; Length 562;  
 Matches 562; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10 CTATCACACACAGTCATGATGATTAACAAACAGTGAAGAGCTCAAGATCTCTGA 69  
 DB 1 CTATCACACACAGTCATGATGATTAACAAACAGTGAAGAGCTCTCAAGATCTCTGA 60  
 QY 70 AGTGGAGAAAGGGCTTGACATGAGAGAGAGCTGATGATGATGATGATGATGATGATG 129  
 DB 61 AGTGGAGAAAGGGCTTGACATGAGAGAGAGCTGATGATGATGATGATGATGATGATG 120  
 QY 130 TCATGGGAGAGGTGTTGGAACCTTTGGCCAAAGCTGCTGCTCAAGCTAAGGAAA 189  
 DB 121 TCATGGGAGAGGTGTTGGAACCTTTGGCCAAAGCTGCTGCTCAAGCTAAGGAAA 180  
 QY 190 GAGTGGCCGCTAAGGTGCTAAATTAATCTCCGCTGCTGATGATGATGATGATGATG 249  
 DB 181 GAGTGGCCGCTAAGGTGCTAAATTAATCTCCGCTGCTGATGATGATGATGATGATG 240  
 QY 250 ACCGAGAGACACTTTGATTTATGAGAGCTCCAGCCAAAGTGGGAAAAGAGTGTCCAA 309  
 DB 241 ACCGAGAGACACTTTGATTTATGAGAGCTCCAGCCAAAGTGGGAAAAGAGTGTCCAA 300  
 QY 310 AATGCGAAGCATCTACTGGAAGAGCTGATGATGATGATGATGATGATGATGATGATG 369  
 DB 301 AATGCGAAGCATCTACTGGAAGAGCTGATGATGATGATGATGATGATGATGATGATG 360  
 QY 370 GATCCAGAGACATCAAGCAAGCTGAGAGACTTTCAGCAACAGAGTAGTATATTCGA 429  
 DB 361 GATCCAGAGACATCAAGCAAGCTGAGAGACTTTCAGCAACAGAGTAGTATATTCGA 420  
 QY 430 GATAATATATCAACAGCTAGCAAGCTGATGATGATGATGATGATGATGATGATGATG 489  
 DB 421 GATAATATATCAACAGCTAGCAAGCTGATGATGATGATGATGATGATGATGATGATG 480  
 QY 490 GATATCTCCACCTGTATCAAGAGATGTAGAGCATTTCACTCACTGCTCTACAT 549  
 DB 481 GATATCTCCACCTGTATCAAGAGATGTAGAGCATTTCACTCACTGCTCTACAT 540  
 QY 550 TAATCCTGATCAATCCAGTTGT 571  
 |||||||

DB 541 TAATCCTGATCAATCCAGTTGT 562

## RESULT 2

BE820766/c

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Glycine.

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

1 (bases 1 to 586)

Vodkin, L., Keim, P., Shoemaker, R., Retzel, E., Khanna, A., Corryell, V.,

Erpelding, J., Rapp, C., Shoop, E., Pardinas, D., Liu, L., and Lewin, H.

A Functional Genomics Program for Soybean (NSF 9872565)

Unpublished (1999)

Other ESTs: AW459279 corresponding to Gm-cl016-5560 (5')

Contact: Vodkin, L.O., PI, A Functional Genomics Program for

Soybean (NSF 9872565)

Lewin, H. A., Director, Keck Center for Comparative and Functional

Genomics

University of Illinois

Edwin R. Madigan Building, 1201 W. Gregory, Urbana, IL 61801, USA

Tel: (217) 244-6147

Fax: (217) 333-4582

Email: l-vodkin@uiuc.edu

This clone is available through: Genome Systems, Inc. 4633 World

Parway Circle St. Louis, Missouri 63134. For further information

call: (800) 450-0030 or (314) 427-3222 FAX: (888) 919-3324 or (314)

427-3324 or contact: clones@genomesystems.com or info@genome

systems.com web site: www.genomesystems.com

Seq primer: 5'-TTTTTTTTTTTTTTT(A/C/G)-3'.

Location/Qualifiers

1. 586

/organism="Glycine max"

/db\_xref="taxon:3847"

/clone="Gm-r1070-4966"

/clone\_1lb="Gm-r1070"

/note="The library Gm-r1070 is a sequence-driven, rerecked

set of 9,216 clones selected from cDNA libraries from

various tissues and stages of development of soybean that

represent 2,639 sequences from immature cotyledons, 1,770

from immature seed coats, 3,938 from flowers, and 869

from young pods. The 5' ESTs of the source clones from

the different libraries was used to select singletons, or

a representative of each contig, which were rerecked to

form library Gm-r1070. The cDNA clones of the rerecked

Gm-r1070 library were then sequenced at the 3' end. The

contig analysis to select unique genes was performed by

the laboratory of Ernest Retzel, Center for Computational

Genomics and Bioinformatics, University of Minnesota,

http://www.cbc.umn.edu/ResearchProjects/soybean/index.html

Rerecking was performed by Genome Systems, St. Louis,

http://www.genomesystems.com, and 3' sequencing by the

Keck Center for Comparative and Functional Genomics,

University of Illinois,

http://www.life.uiuc.edu/biotech/keck.html. Note: The

corresponding 5' EST from each clone in the Gm-r1070

library is listed in the 'OTHER EST' field. The detailed

information on the source library for each clone can also

be obtained by referring to the Genome Systems clone ID of

'OTHER EST'."

BASE COUNT

ORIGIN

173 a 110 c 100 g 194 t 9 others

## Query Match

58.7%; Score 506.4; DB 12; Length 586;

Best Local Similarity 98.6%; Pred. No. 4.7e-66;  
Matches 507; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 326 CTTGAGAGAGCTGATTAATGAGATCAAGAACTATTGGAGACAGAGATCCAGACACATC 385  
Db 583 CCGAGAGAGCTGATTAATGAGATCAAGAACTATTGGAGACAGAGATCCAGACACATC 524  
QY 386 AAGCAAGCTGAGAACTTTGACACACAGAGTGTATTAATTTCTGAGATTAATGATCCAA 445  
Db 523 AAGCAAGCTGAGAACTTTGACACACAGAGTGTATTAATTTCTGAGATTAATGATCCAA 464  
QY 446 GCTGACATACCAAGTTTCCACCATGCGCCAGCCAGAGATGATATTCACACCTGT 505  
Db 463 GCTGACATACCAAGTTTCCACCATGCGCCAGAGATGATATTCACACCTGT 404  
QY 506 TATCAAGAGATGTTAGAGCCATTTTCAACTGTTCCCTCAATTAATCCGATCAATCC 565  
Db 403 TATCAAGAGATGTTAGAGCCATTTTCAACTGTTCCCTCAATTAATCCGATCAATCC 344  
QY 566 AGTTGTTTACCATGACACACACATTAATTAATTTGAGAGATGAGAGATGCTGTCA 625  
Db 343 AGTTGTTTACCATGACACACACATTAATTAATTTGAGAGATGAGAGATGCTGTCA 284  
QY 626 ATGCAATTTACTGACAGGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 685  
Db 283 ATGCAATTTACTGACAGGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 224  
QY 686 TAAGCTGAGATGTTCTGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 745  
Db 223 TAAGCTGAGATGTTCTGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 164  
QY 746 ATGTTGTTTGGTTCACCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 805  
Db 163 ATGTTGTTTGGTTCACCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 104  
QY 806 AAATATATCTGATCTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 839  
Db 103 AAATATATCTGATCTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 70

RESULT 3  
BE658316/c 782 bp mRNA linear EST 24-MAY-2001  
LOCUS GM700005B10B4 Gm-r1070 Glycine max cDNA clone Gm-r1070-1759 3',  
DEFINITION mRNA sequence.  
ACCESSION BE658316  
VERSION BE658316.1 GI:9984208  
KEYWORDS EST.  
SOURCE soybean.  
ORGANISM Glycine max  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;  
Glycine.  
1 (bases 1 to 782)  
Vodkin,L., Keim,P., Shoemaker,R., Retzel,E., Khanna,A., Coryell,V.,  
Expelling,J., Rapp,C., Shoop,E., Pardinas,J., Liu,L. and Lewin,H.,  
A Functional Genomics Program for Soybean (NSF 9872565)  
Unpublished (1999)  
Other\_ESTS: A1930997 corresponding to Gm-cl015-278 (5')  
Contact: Vodkin, L.O., PI, A Functional Genomics Program for  
Soybean (NSF 9872565)  
Lewin, H. A., Director, Keck Center for Comparative and Functional  
Genomics  
University of Illinois  
Edwin R. Madigan Building, 1201 W. Gregory, Urbana, IL 61801, USA  
Tel: (217) 244-6147  
Fax: (217) 333-4582  
Email: l-vodkin@uiuc.edu

This clone is available through: Genome Systems, Inc. 4633 World  
Parkway Circle St. Louis, Missouri 63134. For further information  
call: (800) 430-0030 or (314) 427-3222 Fax: (888) 919-3324 or (314)  
427-3324 or contact: clones@genomesystems.com or info@genome

systems.com web site: www.genomesystems.com  
Seq primer: 5'-TTTTTTTTTTTTTTTTT(A/C/G)-3'.  
FEATURES  
source  
Location/Qualifiers  
1..782  
/organism="Glycine max"  
/db\_xref="taxon:3847"  
/clone="Gm-r1070-1759"  
/clone\_lib="Gm-r1070"  
/note="The library Gm-r1070 is a sequence-driven, reracked  
set of 9,216 clones selected from cDNA libraries from  
various tissues and stages of development of soybean that  
represent 2,639 sequences from immature cotyledons, 1,770  
from immature seed coats, 3,938 from flowers, and 869  
from young pods. The 5' ESTs of the source clones from  
the different libraries was used to select singletons, or  
a representative of each contig, which were reracked to  
form library Gm-r1070. The cDNA clones of the reracked  
Gm-r1070 library were then sequenced at the 3' end. The  
contig analysis to select unique genes was performed by  
the laboratory of Ernest Retzel, Center for Computational  
Genomics and Bioinformatics, University of Minnesota.  
http://www.cbc.umn.edu/Research/Projects/Soybean/Index.html  
Reracking was performed by Genome Systems, St. Louis,  
http://www.genomesystems.com, and 3' sequencing by the  
Keck Center for Comparative and Functional Genomics,  
University of Illinois,  
http://www.life.uiuc.edu/biotech/keck.html. Note: The  
corresponding 5' EST from each clone in the Gm-r1070  
library is listed in the 'OTHER EST' field. The detailed  
information on the source library for each clone can also  
be obtained by referring to the Genome Systems clone ID of  
the original cDNA library that is also listed under  
'OTHER EST'."

BASE COUNT 215 a 144 c 143 g 253 t 27 others  
ORIGIN

Query Match 56.9%; Score 491.4; DB 10; Length 782;  
Best Local Similarity 89.3%; Pred. No. 6.5e-64;  
Matches 612; Conservative 2; Mismatches 49; Indels 22; Gaps 8;

QY 111 TGATGAAGTAAATGCAATCATGGGAGAGTGTGGAACTTTGGCCAAAGCTGCTG 170  
Db 782 TGATGAAGTAAATGCAATCATGGGAGAGTGTGGAACTTTGGCCAAAGCTGCTG 723  
QY 171 GGTCAAAAGTAAAGGAGAGTGGCGGCTAAGGTGCTAAATTAATTAATTAATTAATTA 230  
Db 722 NCTCAAAAGTAAAGGAGAGTGGCGGCTAAGGTGCTAAATTAATTAATTAATTAATTA 663  
QY 231 TTGAAGAGGGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 290  
Db 662 TTGAAGAGGGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 603  
QY 291 GGGGAAGAGGTGGTCCAAATTTGGCAAGCATCTCCGTGGAAGCATGATATATGATCA 350  
Db 602 GGGGAAGAGGTGGTCCAAATTTGGCAAGCATCTCCGTGGAAGCATGATATATGATCA 543  
QY 351 AGAATATTGGAGGACAMAGATCCAGAACACATCAAGCATCAAGCATCAAGCATCAAG 410  
Db 542 AGAATATTGGAGGACAMAGATCCAGAACACATCAAGCATCAAGCATCAAGCATCAAG 483  
QY 411 AAGAGTAAATTAATTTCTGAGTAATTAATTAATTAATTAATTAATTAATTAATTAAT 470  
Db 482 AAGAGTAAATTAATTTCTGAGTAATTAATTAATTAATTAATTAATTAATTAATTAAT 423  
QY 471 TGGCTAGGCCATGAGAGATGTTCTCCACCTGTTATCAAGAGATGTTAAGCATTT 528  
Db 422 TGGCTAGGCCATGAGAGATGTTCTCCACCTGTTATCAAGAGATGTTAAGCATTT 363  
QY 529 -TTCAACTAGTTCCCTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 587  
Db 362 CTTCAATTAAGTTCCCTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 303  
QY 588 ACAA---CATTAACATTTGGAGCATGAGAGATAGCTGTCAATG-CAATTAAGTCAAGCGT 643

```

Db 302 ACACACGATTAAGTATGAGCATGAGATATCTGCTAAATGATAGTACGAAACGGG 243
QY 644 GATTAAATAT---TATCAAGATAAAACCTAAGTTTG---AAGTTCAT--AAGCTG 693
Db 242 GATTAAATATGATATATCAAGATAAAACCTAATCTGTTATATAGTCCATAAACCTG 183
QY 694 GAATGCTGTGGATTAACA---TATTATGGTGTGTTATATAGTAGTGG--ATG 748
Db 182 GAATGCTGTGGCTTAACAATATATATAGTGTGTTATATATAGTAGTGGATATG 123
QY 749 TTGGTTTGGGATACATTAATAGC 773
Db 122 TTGGTTTGGGATACATTAATAGC 98

RESULT 4
BM527606 501 bp mRNA linear EST 19-FEB-2002
LOCUS BM527606
DEFINITION BM527606 y1 Gm-cl061 Glycine max cDNA clone SOYBEAN CLONE ID:
sequence. Gm-cl061-3780 5' similar to YR:039028 Q39028 ATMB2.1; mRNA
sequence.
ACCESSION BM527606.1 GI:18733434
VERSION BM527606
KEYWORDS EST.
SOURCE soybean.
ORGANISM Glycine max
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
Glycine.
REFERENCE 1 (bases 1 to 501)
AUTHORS Shoemaker,R., Keim,P., Vodkin,L., Erpelting,J., Coryell,V., Khanna
A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C.,
Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers
Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk
R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann
R., Waterson,R. and Wilson,R.
TITLE Public Soybean EST Project
JOURNAL Unpublished (1999)
COMMENT Contact: Shoemaker R/Public Soybean EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available through: Resgen, Invitrogen Corp. 2130
South Memorial Parkway Huntsville, AL 35801 For further information
call: (800)-533-4363 or contact: ccl@resgen.com web site:
www.resgen.com
Seq primer: -40RP from GIBCO
High quality sequence stop: 421.
Location/Qualifiers
1. 501
FEATURES
SOURCE
/organism="Glycine max"
/db_xref="taxon:847"
/clone="SOYBEAN CLONE ID: Gm-cl061-3780"
/clone_id="Gm-cl061"
/lisuse_type="mature flowers of field grown plants"
/lab_host="DH10B"
/notes="Vector: plusscript II SK+; Site.1: EcoRI; Site.2:
XhoI; The cDNA library was constructed from mRNA isolated
from mature flowers of field grown plants for the cultivar
Raiden. Complementary DNA was synthesized from mRNA using
a primer consisting of a poly(dT) sequence ligated to the
restriction site. EcoRI adapters followed by XhoI digestion. The
cDNA fragments were directionally cloned into the
EcoRI-XhoI restriction site of the pluscript vector. The
ligated cDNA fragments were transformed into DH10B host
cells (GibcoBRL). This library was constructed in the
laboratory of Dr. Randy Shoemaker."

```

```

BASE COUNT 170 a 104 c 114 g 113 t
ORIGIN

```

```

Query Match 53.6%; Score 462.6; DB: 13; Length 501;
Best Local Similarity 95.2%; Fred. No. 1.5e-59;
Matches 477; Conservative 0; Mismatches 24; Indels 0; Gaps 0;

```

```

QY 20 CACAAGTCATGATTAACAAAACACAGTGTAAAGCTCAAGATCTGAAGTGAAGAAA 79
Db 1 CACAACCAATGATTAACAAAACACAGTGTAAAGCTCAAGATCTGAAGTGAAGAAA 60
QY 80 GGGCCCTGGCAATGAAGAAAGACTGATGATGATGATGATGATGATGATGATGATG 139
Db 61 GGACCTGGACATGAGAAAGACTGATGATGATGATGATGATGATGATGATGATGATG 120
QY 140 GGATGTTGAACTCTTTGGCCAAAGCTGCTCTCAAAAGCTGAAGAAAGATGGCCGG 199
Db 121 GGATGTTGAAATCTTTGGCCAAAGCTGCTCTCAAAAGCTGAAGAAAGATGGCCGG 180
QY 200 CTAAGGTGGCTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 259
Db 181 CTAAGGTGGCTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 240
QY 260 CAACCTTTGATTAATGAGCTCCAGCAAGAGTGGGAAACAGGTGCTCAAAATTCGCAAG 319
Db 241 CAACCTTTGATTAATGAGCTCCAGCAAGAGTGGGAAACAGGTGCTCAAAATTCGCAAG 300
QY 320 CATCTACCTGGAGAGAGCTGATATGATGATGATGATGATGATGATGATGATGATGAT 379
Db 301 CATCTACCTGGAGAGAGCTGATATGATGATGATGATGATGATGATGATGATGATGAT 360
QY 380 CACATCAACCAAGCTGAGAGACTTTCAGCAGAGTAGTATTAATTCGATTAATGAT 439
Db 361 CACATCAACCAAGCTGAGAGACTTTCAGCAGAGTAGTATTAATTCGATTAATGAT 420
QY 440 CACCAAGTAGCAGTACGATGATGATGATGATGATGATGATGATGATGATGATGAT 499
Db 421 CACCAAGTAGCAGTACGATGATGATGATGATGATGATGATGATGATGATGATGAT 480
QY 500 CCCTGTTATCAAGAAATGTTA 520
Db 481 CCCTGTTATCAAGAAATGTTA 501

```

```

RESULT 5
BE058947 450 bp mRNA linear EST 03-DEC-2001
LOCUS BE058947
DEFINITION sn23b01 y1 Gm-cl016 Glycine max cDNA clone GENOME SYSTEMS CLONE ID:
Gm-cl016-12290 5' similar to YR:027D5 Q927D5 PUTATIVE
TRANSCRIPTION FACTOR ;, mRNA sequence.
ACCESSION BE058947 GI:8403313
VERSION BE058947
KEYWORDS EST.
SOURCE soybean.
ORGANISM Glycine max
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
Glycine.
REFERENCE 1 (bases 1 to 450)
AUTHORS Shoemaker,R., Keim,P., Vodkin,L., Erpelting,J., Coryell,V., Khanna
A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C.,
Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers
Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk
R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann
R., Waterson,R. and Wilson,R.
TITLE Public Soybean EST Project
JOURNAL Unpublished (1999)
COMMENT Contact: Shoemaker R/Public Soybean EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800

```

Fax: 314 286 1810  
Email: estewatson.wustl.edu  
This clone is available through: ResGen, Invitrogen Corp. 2130 South Memorial Parkway Huntville, AL 35801 For further information call: (800)-533-4363 or contact via email: ccu@resgen.com  
Seq primer: -40RP from Gibco  
High quality sequence stop: 411.

## FEATURES

## source

Location/Qualifiers  
1..450  
/organism="Glycine max"  
/db\_xref="taxon:3847"  
/clone="GENOME SYSTEMS CLONE ID: Gm-c1016-12290"  
/clone\_1ib="Gm-c1016"  
/tissue\_type="Immature flowers of field grown plants"  
/lab\_host="X110-Gold"  
/note="Vector: pBluescript II XR; Site\_1: EcoRI; Site\_2: XhoI; This cDNA library was constructed from mRNA isolated from immature flowers of field grown plants. The cDNA library was prepared using the Stratagene pBluescript II XR library construction kit. Complementary DNA was synthesized from mRNA using a primer consisting of a poly (dT) sequence with a XhoI restriction site. EcoRI adapters were ligated to the blunt-ended cDNA fragments followed by XhoI digestion. The cDNA fragments were directionally cloned into the EcoRI-XhoI restriction site of the pBluescript vector. The ligated cDNA fragments were transformed into X110-Gold host cells. This library was constructed by Dr. Randy Shoemaker and Dr. John Erpelidg."

BASE COUNT 151 a 84 c 90 g 125 t  
ORIGIN

## Query Match

Best Local Similarity 98.9%; Score 444.4; DB 10; Length 450;  
Matches 445; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 304 GTCCAAAATGCCAGATCTACCTGGAGGAGTGAATGAGATCAAGACATTGGAG 363  
DB 1 GTCCAAAATGCCAGATCTACCTGGAGGAGTGAATGAGATCAAGACATTGGAG 60  
QY 364 GACAAAGATCCAGACACATCAAGCAAGCTGGAATCTTCAGCAAGAGTGAATAA 423  
DB 61 GACAAAGATCCAGACACATCAAGCAAGCTGGAATCTTCAGCAAGAGTGAATAA 120  
QY 424 TTCTGAGATTAATGATCACCAGCTAGCAGTTCACCATGGCTGAGCCAT 483  
DB 121 TTCTGAGATTAATGATCACCAGCTAGCAGTTCACCATGGCTGAGCCAT 180  
QY 484 GGAGATGATTTCCACACCTGTTATCAAGGAATGAGGCCATTTTCAACTGATGCC 543  
DB 181 GGAGATGATTTCCACACCTGTTATCAAGGAATGAGGCCATTTTCAACTGATGCC 240  
QY 544 TACAAATTAATCCTGATCAATCCAGTTGTTGTACCAATGACACACACATTAATG 603  
DB 241 TACAAATTAATCCTGATCAATCCAGTTGTTGTACCAATGACACACATTAATG 300  
QY 604 GAGCAATGAGGATAGCTGGTCAATTCGAAACGGTGAATTAATATATCAAGATA 663  
DB 301 GAGCAATGAGGATAGCTGGTCAATTCGAAACGGTGAATTAATATATCAAGATA 360  
QY 664 AAACCTAAGTTTGAAGTTCCATAGCGTAATGCTTGATTAACATATTTTGG 723  
DB 361 AAACCTAAGTTTGAAGTTCCATAGCGTGAATGCTTGATTAACATATTTTGG 420  
QY 724 GTTGTGTTATATAAGTATGATGTTGG 753  
DB 421 GTTGTGTTATATAAGTATGATGTTGG 450

RESULT 6  
LOCUS A1930997 523 bp mRNA linear EST 30-NOV-2001  
DEFINITION sb45h07.y1 Gm-c1015 Glycine max cDNA clone GENOME SYSTEMS CLONE ID:

ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM

## REFERENCE

## AUTHORS

## JOURNAL

1 (bases 1 to 523)  
Shoemaker, R., Keim, P., Vodkin, L., Erpelidg, J., Coryell, V., Khanna, A., Bolla, B., Merritt, M., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wylie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers, Y., Person, B., Swaller, T., Gibbons, M., Page, D., Harvey, N., Schurk, R., Rittner, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R., Waterston, R. and Wilson, R.  
Public Soybean EST Project  
Unpublished (1999)  
Contact: Shoemaker R/Public Soybean EST Project  
Public Soybean EST Project  
Washington University School of Medicine  
444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: estewatson.wustl.edu  
This clone is available through: ResGen, Invitrogen Corp. 2130 South Memorial Parkway Huntville, AL 35801 For further information call: (800)-533-4363 or contact via email: ccu@resgen.com  
Seq primer: -40RP from Gibco  
High quality sequence stop: 422.

## FEATURES

## source

Location/Qualifiers  
1..523  
/organism="Glycine max"  
/db\_xref="taxon:3847"  
/clone="GENOME SYSTEMS CLONE ID: Gm-c1015-278"  
/clone\_1ib="Gm-c1015"  
/tissue\_type="Mature flowers, field grown plants"  
/lab\_host="X110-Gold"  
/note="Vector: pBluescript II XR; Site\_1: EcoRI; Site\_2: XhoI; This cDNA library was constructed from mRNA isolated from mature flowers of field grown plants. The cDNA library was prepared using the Stratagene pBluescript II XR cDNA library construction kit. Complementary DNA was synthesized from mRNA using a primer consisting of a poly (dT) sequence with a XhoI restriction site. EcoRI adapters were ligated to the blunt-ended cDNA fragments followed by XhoI digestion. The cDNA fragments were directionally cloned into the EcoRI-XhoI restriction site of the pBluescript vector. The ligated cDNA fragments were transformed into X110-Gold host cells. This library was constructed by Dr. Randy Shoemaker and Dr. John Erpelidg."

BASE COUNT 172 a 124 c 111 g 115 t 1 others  
ORIGIN

## Query Match

Best Local Similarity 94.3%; Score 442.8; DB 9; Length 523;  
Matches 459; Conservative 0; Mismatches 28; Indels 0; Gaps 0;

QY 8 CTCATACACACACAGTCAATGATTAATAAACAACAGTGAAGCTCTCAAGATCCT 67  
DB 37 CTCATACACACACACAAATGATTAATAAACAACAGTGAAGCTCTCAAGATCCT 96  
QY 68 GAAGTGAAGAAAGGCGCTGGATGATGAAGAAGCTGATCTGATGATCAATATTTGCA 127  
DB 97 GAAGTGAAGAAAGGCGCTGGATGATGAAGAAGCTGATCTGATGATCAATATTTGCA 156  
QY 128 AATCATGGGGAAGGCTTTGGAACTTTGGCCAAAGCTGCTGCTCAACGTAAGGA 187  
DB 157 AATCATGGGGAAGGCTTTGGAACTTTGGCCAAAGCTGCTGCTCAACGTAAGGA 216

QY 188 AAGAGTCCGCGCTAAGGTGGCTAAATTACCTCCGCTCTGATGTTAGAGAGGGAATATT 247  
 |||||  
 Db 217 AAGAGTCCGCGCTAAGGTGGCTAAGCTACCTCCGCTCTGATGTTAGAGAGGGAATATT 276  
 |||||  
 QY 248 ACACCCGAGGAACAACCTTTGATATGAGAGCTCCACCAAGTGGGAACAGGTGGTCC 307  
 |||||  
 Db 277 ACACCCGAGGAACAACCTTTGATATGAGAGCTCCACCAAGTGGGAACAGGTGGTCC 336  
 |||||  
 QY 308 AAAATTGCCAAGCATCTACCTGGAAGACATGATATGAGATCAAGAACTATTGGAGAGACA 367  
 |||||  
 Db 337 AAAATTGCCAAGCATCTACCTGGAAGACATGATATGAGATCAAGAACTATTGGAGAGACC 396  
 |||||  
 QY 368 AGGATCCAGGAAGCATCAACAGCTGAGAACTTTCAGCAAGAGATGATTAATTTCT 427  
 |||||  
 Db 397 AGGATCCAGGAAGCATCAACAGCTGAGAACTTTCAGCAAGAGATGATTAATTTCT 456  
 |||||  
 QY 428 GAGATTAATGATCAACAGCTAGGACATGAGCATGTTTCCACCATGGCTGAGCCCATGAG 487  
 |||||  
 Db 457 GAGATTAATGATCAACAGCTAGGACATGAGCATGTTTCTACCATGGCTGAGAACCATGAG 516  
 |||||  
 QY 488 ATGATTT 494  
 |||||  
 Db 517 ACTGATTT 523

RESULT 7  
 AM459279 431 bp mRNA linear EST 03-DEC-2001  
 LOCUS sh22h08.y1 Gm-c1016 Glycine max cDNA clone GENOME SYSTEMS CLONE ID:  
 DEFINITION Gm-c1016-5560 5' similar to TR:09ZTD5 09ZTD5 PUTATIVE TRANSCRIPTION  
 FACTOR ; mRNA sequence.

ACCESSION AM459279  
 VERSION AM459279.1 GI:7029496  
 KEYWORDS EST.

SOURCE soybean.  
 ORGANISM Glycine max

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;  
 Glycine.

REFERENCE 1 (bases 1 to 431)  
 AUTHORS Shoemaker,R., Keim,P., Vodkin,L., Erpelting,J., Corvelli,V., Khanna  
 A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C.,  
 Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers  
 Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk  
 R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann  
 R., Waterston,R. and Wilson,R.

TITLE Public Soybean EST Project  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Shoemaker R/Public Soybean EST Project  
 Public Soybean EST Project  
 Washington University School of Medicine  
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA  
 Tel: 314 286 1800  
 Fax: 314 286 1810  
 Email: est@wustl.wustl.edu

This clone is available through: ResGen, Invitrogen Corp. 2130  
 South Memorial Parkway Huntville, AL 35801 For further information  
 call: (800)-533-4363 or contact via email: ccu@resgen.com  
 Insert length: 723 Std Error: 0.00  
 Seq primer: -40RP from Gibco  
 High quality sequence stop: 411.

FEATURES  
 SOURCE

Location/Qualifiers  
 1..431  
 /organism="Glycine max"  
 /db\_xref="taxon:3847"  
 /clone="GENOME SYSTEMS CLONE ID: Gm-c1016-5560"  
 /clone\_id="Gm-c1016"  
 /tissue\_type="immature flowers of field grown plants"  
 /lab\_host="XL10-GOLD"  
 /note="Vector: plusscript II XR; Site\_1: EcoRI; Site\_2:  
 XhoI; This cDNA library was constructed from mRNA isolated  
 from immature flowers of field grown plants. The cDNA

library was prepared using the Stratagene pluscript II  
 XR library construction kit. Complementary DNA was  
 synthesized from mRNA using a primer consisting of a poly  
 (dT) sequence with a XhoI restriction site. EcoRI adapters  
 were ligated to the blunt-ended cDNA fragments followed by  
 XhoI digestion. The cDNA fragments were directionally  
 cloned into the EcoRI-XhoI restriction site of the  
 pluscript vector. The ligated cDNA fragments were  
 transformed into XL10-Gold host cells. This library was  
 constructed by Dr. Randy Shoemaker and Dr. John  
 Erpelting."

BASE COUNT 145 a 81 c 88 g 117 t  
 ORIGIN  
 Query Match 49.3%; Score 425.4; DB 10; Length 431;  
 Best Local Similarity 98.8%; Pred. No. 5.4e-54;  
 Matches 426; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 319 GCATCTACCTGGAAGCATGATATGATCAAGAACTATTGGAGAGAGATGATTTCC 378  
 |||  
 Db 1 GCACACGCTCGAAGCATGATATGATCAAGAACTATTGGAGAGAGATGATTTCC 60  
 |||||  
 QY 379 GCACATCAAGCAAGCTGAGAACTTTTCAGCAAGAGATGATTAATTTTCAGATTAATGA 438  
 |||||  
 Db 61 GCACATCAAGCAAGCTGAGAACTTTTCAGCAAGAGATGATTAATTTTCAGATTAATGA 120  
 |||||  
 QY 439 TCACCAAGCTAGACATGAGCATGTTTCCACCATGGCTGAGCCCATGAGATGATTTCC 498  
 |||||  
 Db 121 TCACCAAGCTAGACATGAGCATGTTTCCACCATGGCTGAGCCCATGAGATGATTTCC 180  
 |||||  
 QY 499 ACCCTGTTATCAAGAAATGTGAGAGCATTTTCAACTGATTCCTACATTAATTCGGA 558  
 |||||  
 Db 181 ACCCTGTTATCAAGAAATGTGAGAGCATTTTCAACTGATTCCTACATTAATTCGGA 240  
 |||||  
 QY 559 TCAATCAGTGTGTGACCAATGACAAACAACTTAATTAATGAGATGAGAGATGAG 618  
 |||||  
 Db 241 TCAATCAGTGTGTGACCAATGACAAACAACTTAATTAATGAGATGAGAGATGAG 300  
 |||||  
 QY 619 CTGCTCAATGCAATTTCTGAGCGTGATTAATTTATCAAGATAAACCCTAAGTTTGA 678  
 |||||  
 Db 301 CTGCTCAATGCAATTTCTGAGCGTGATTAATTTATCAAGATAAACCCTAAGTTTGA 360  
 |||||  
 QY 679 AGTTCATTAAGGCTGGAAGTCTCTTGATTAACATATTATGAGTTGTTATATAG 738  
 |||||  
 Db 361 AGTTCATTAAGGCTGGAAGTCTCTTGATTAACATATTATGAGTTGTTATATAG 420  
 |||||  
 QY 739 TACTTGGATGT 749  
 |||||  
 Db 421 TACTTGGATGT 431

RESULT 8  
 BM527774 552 bp mRNA linear EST 19-FEB-2002  
 LOCUS sal55903.y1 Gm-c1061 Glycine max cDNA clone SOYBEAN CLONE ID:  
 DEFINITION Gm-c1061-4134 5' similar to TR:049020 049020 MYB-LIKE DNA-BINDING  
 DOMAIN PROTEIN. ; mRNA sequence.

ACCESSION BM527774  
 VERSION BM527774.1 GI:18733722  
 KEYWORDS EST.

SOURCE soybean.  
 ORGANISM Glycine max

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;  
 Glycine.

REFERENCE 1 (bases 1 to 552)  
 AUTHORS Shoemaker,R., Keim,P., Vodkin,L., Erpelting,J., Corvelli,V., Khanna  
 A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C.,  
 Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers  
 Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk  
 R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann  
 R., Waterston,R. and Wilson,R.

TITLE  
JOURNAL  
COMMENT

Public Soybean EST Project  
Unpublished (1999)  
Contact: Shoemaker R/Public soybean EST Project  
Public Soybean EST Project  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: est@wustl.edu  
This clone is available through: Resgen, Invitrogen Corp. 2130  
South Memorial Parkway Huntsville, AL 35801 For further information  
call: (800)-533-4363 or contact: c@resgen.com web site:  
www.resgen.com  
Seq primer: -40RP from Gibco  
High quality sequence stop: 426.  
Location/Qualifiers  
1. .552  
/organism="Glycine max"  
/db\_xref="taxon:3847"  
/clone="SOYBEAN CLONE ID: Gm-cl061-4134"  
/clone\_lib="Gm-cl061"  
/tissue\_type="mature flowers of field grown plants"  
/lab\_host="DH10B"  
/note="Vector: pBluescript II SK+; Site.1: EcoRI; Site.2:  
XhoI; The cDNA library was constructed from mRNA isolated  
from mature flowers of field grown plants for the cultivar  
Raiden. Complementary DNA was synthesized from mRNA using  
a primer consisting of a poly(dT) sequence with a XhoI  
restriction site. EcoRI adapters were ligated to the  
blunt-ended cDNA fragments followed by XhoI digestion. The  
cDNA fragments were directionally cloned into the  
EcoRI-XhoI restriction site of the pBluescript vector. The  
ligated cDNA fragments were transformed into DH10B host  
cells (GibcoBRL). This library was constructed in the  
laboratory of Dr. Randy Shoemaker."

## FEATURES

source

BASE COUNT 184 a 118 c 111 g 139 t  
ORIGIN

Query Match 49.2%; Score 424.4; DB 13; Length 552;  
Best Local Similarity 91.7%; Pred. No. 6.5e-54;

Matches 506; Conservative 2; Mismatches 28; Indels 16; Gaps 5;

OY 173 CTCGAACGTACGGAAGAGTCCCGCTAAGTGGCTAAATTACCTCCGCTGATGTT 232  
Db 1 CTCGAACGTACGGAAGAGTGGCGGCTAAGTGGCTAACTACCTCCGCTGATGTT 60  
OY 233 AGAAGAGGGAATATTCACCCGAGAACACTTTGATTTAGAGCTCCAGCAAAATGG 292  
Db 61 AGAAGAGGGAATATTCACCCGAGAACACTTTGATTTAGAGCTCCAGCAAAATGG 120  
OY 293 GGAACAGGTGGTCCAAATTCGCAAGCATCTACTGGAAGAGCATGATATGATCAAG 352  
Db 121 GGAACAGGTGGTCCAAATTCGCAAGCATCTACTGGAAGAGCATGATATGATCAAG 180  
OY 353 AACTATTGAGAGACAGATCCAGAACATCAAGCAAGTGAAGTTCACACAG 412  
Db 181 AACTATTGAGAGACAGATCCAGAACATCAAGCAAGTGAAGTTCACACAG 240  
OY 413 AGTAGTAATATTCGAGATTAATGATACCAAGCTAGACACTAGCTGTTTCCACATG 472  
Db 241 AGTAGTAATATTCGAGATTAATGATACCAAGCTAGACACTAGCTGTTTCCACATG 300  
OY 473 GGTGAGCCCATGAGATGATTCGACCCCTGTATCAAGGAATGTAGAGCAT---T 529  
Db 301 GGTGAGCCCATGAGATGATTCGACCCCTGTATCAAGGAATGTAGAGCATTTTCT 360  
OY 530 TCAACTGACCTCCCTACAAATTAATCCGATCAATCAGTTGTTTACCAATGACAAC 589  
Db 361 TCAATTCAGTCCCAATTAATCCGATCAATCAGTTGTTTACCAATGACAAC 420  
OY 590 AA---CATTAATATTTGAGAGCATGAGATAGCTGGTCAATGCAATTTACTGAGCGTAT 646  
Db 421 AACGACATTAACTATTGAGAGCATGAGATATCTGTCATATCCAGTTACTGACGCGGAT 480

OY 647 TAAATAT-----TATCAAGATTAACCTAAGTTTGG---AAGTCCAT--AAGCTGGA 696  
Db 481 TAAATATGATATATCAAGATTAACCTAAGTTTGTATAGTTCACAAACACTGGA 540  
OY 697 TGTCTGTGATG 708  
Db 541 TGTCTGTGCTT 552

## RESULT 9

LOCUS

DEFINITION

BM527664 568 bp mRNA linear EST 19-FEB-2002  
sa164610.y1 Gm-cl061 Glycine max cDNA clone SOYBEAN CLONE ID:  
Gm-cl061-4003 5' similar to TR:Q39028 Q39028 ATMB2. ;, mRNA  
sequence.

## ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Public Soybean EST Project  
Unpublished (1999)  
Contact: Shoemaker R/Public soybean EST Project  
Public Soybean EST Project  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: est@wustl.edu  
This clone is available through: Resgen, Invitrogen Corp. 2130  
South Memorial Parkway Huntsville, AL 35801 For further information  
call: (800)-533-4363 or contact: c@resgen.com web site:  
www.resgen.com  
Seq primer: -40RP from Gibco  
High quality sequence stop: 422.  
Location/Qualifiers  
1. .568  
/organism="Glycine max"  
/db\_xref="taxon:3847"  
/clone="SOYBEAN CLONE ID: Gm-cl061-4003"  
/clone\_lib="Gm-cl061"  
/tissue\_type="mature flowers of field grown plants"  
/lab\_host="DH10B"  
/note="Vector: pBluescript II SK+; Site.1: EcoRI; Site.2:  
XhoI; The cDNA library was constructed from mRNA isolated  
from mature flowers of field grown plants for the cultivar  
Raiden. Complementary DNA was synthesized from mRNA using  
a primer consisting of a poly(dT) sequence with a XhoI  
restriction site. EcoRI adapters were ligated to the  
blunt-ended cDNA fragments followed by XhoI digestion. The  
cDNA fragments were directionally cloned into the  
EcoRI-XhoI restriction site of the pBluescript vector. The  
ligated cDNA fragments were transformed into DH10B host  
cells (GibcoBRL). This library was constructed in the  
laboratory of Dr. Randy Shoemaker."

## FEATURES

source

BASE COUNT 174 a 134 c 124 g 136 t  
ORIGIN

Query Match 46.5%; Score 401.2; DB 13; Length 568;  
Best Local Similarity 85.6%; Pred. No. 1.7e-50;  
Matches 471; Conservative 0; Mismatches 73; Indels 6; Gaps 2;





```

DEFINITION      ss45f11.y1 Gm-c1061 Glycine max cDNA clone GENOME SYSTEMS CLONE ID:
                  Gm-c1061-1702 5' similar to TR:039028 Q39028 ATMYB2. ;, mRNA
SEQUENCE
ACCESSION      BE804790
VERSION        BE804790.1 GI:10235902
KEYWORDS
SOURCE         soybean.
ORGANISM       Glycine max
                Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
                Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
                Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
                Glycine
REFERENCE
AUTHORS        1 (bases 1 to 409)
                Shoemaker,R., Keim,P., Vodkin,L., Erpelting,J., Coryell,V., Khanna
                ,A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C.,
                Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers
                ,Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk
                ,R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann
                ,R., Waterston,R. and Wilson,R.
                Public Soybean EST Project
                Unpublished (1999)
JOURNAL
COMMENT        Contact: Shoemaker R/Public Soybean EST Project
                Washington University School of Medicine
                4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
                Tel: 314 286 1800
                Fax: 314 286 1810
                Email: est@watson.wustl.edu
                This clone is available through: Resgen, Invitrogen Corp. 2130
                South Memorial Parkway Huntville, AL 35801 For further information
                call: (800)-533-4363 or contact via email: ccu@resgen.com
                Insert Length: 1068 Std Error: 0.00.
FEATURES
source
1..409
/organism="Glycine max"
/db_xref="taxon:3847"
/clone="GENOME SYSTEMS CLONE ID: Gm-c1061-1702"
/tissue_type="mature flowers of field grown plants"
/lab_host="DH10B"
/note="Vector: pBluescript II SK+; Site:1: EcoRI; Site:2:
XhoI; The cDNA library was constructed from mRNA isolated
from mature flowers of field grown plants for the cultivar
Raiden. Complementary DNA was synthesized from mRNA using
a primer consisting of a poly(dT) sequence with a XhoI
restriction site. EcoRI adapters were ligated to the
blunt-ended cDNA fragments followed by XhoI digestion. The
cDNA fragments were directionally cloned into the
EcoRI-XhoI restriction site of the pBluescript vector. The
ligated cDNA fragments were transformed into DH10B host
cells (GibcoBRL). This library was constructed in the
laboratory of Dr. Randy Shoemaker."
BASE COUNT      136 a      87 c      97 g      89 t
ORIGIN
Query Match      42.9%; Score 369.8; DB 12; Length 409;
Best Local Similarity 95.7%; Pred. No. 9.4e-6;
Matches 380; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

```

```

Db 193 AACAGTCCGCGGCTAAGTGCGCTAAACTACCTCCGCTCGATGATTGAAGAGCGGAATATT 252
Qy 248 ACACCCGAGAACCAACTTTTGAATTATGAGCTCCAGCAAGTGGGAAACAGTGTC 307
Db 253 ACACCCGAGAACCAACTTTTGAATTATGAGCTTCACGCAAAATGGGAAACAGTGTC 312
Qy 308 AAAATTGCCAAGCATCTACCTGGAAAGCTGATTAATGAGATCAAGAAGTAATTGGAGACA 367
Db 313 AAAATTGCCAAGCATCTACCTGGTGGACAGATTAATGAGATCAAGATTAATTGGAGACC 372
Qy 368 AGGATCCAGAGACATCAAGCAAGCTGAGAACTTTC 404
Db 373 AGGATCCAGAGACATCAAGCAAGCTGAGAACTTTC 409
RESULT 12
LOCUS          BM732539
DEFINITION     BM732539 560 bp mRNA linear EST 01-MAR-2002
                sal78e08.y1 Gm-c1061 Glycine max cDNA clone SOYBEAN CLONE ID:
                Gm-c1061-5223 5' similar to TR:Q39028 Q39028 ATMYB2. ;, mRNA
                sequence.
ACCESSION      BM732539
VERSION        BM732539
KEYWORDS
SOURCE         EST.
ORGANISM       soybean.
                Glycine max
                Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
                Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
                Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
                Glycine
                1 (bases 1 to 560)
REFERENCE
AUTHORS        Shoemaker,R., Keim,P., Vodkin,L., Erpelting,J., Coryell,V., Khanna
                ,A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C.,
                Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers
                ,Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk
                ,R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann
                ,R., Waterston,R. and Wilson,R.
                Public Soybean EST Project
                Unpublished (1999)
JOURNAL
COMMENT        Contact: Shoemaker R/Public Soybean EST Project
                Washington University School of Medicine
                4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
                Tel: 314 286 1800
                Fax: 314 286 1810
                Email: est@watson.wustl.edu
                This clone is available through: Resgen, Invitrogen Corp. 2130
                South Memorial Parkway Huntville, AL 35801 For further information
                call: (800)-533-4363 or contact: ccu@resgen.com web site:
                www.resgen.com
                Seq primer: -40RP from Gibco
                High quality sequence stop: 430.
FEATURES
source
1..560
/organism="Glycine max"
/db_xref="taxon:3847"
/clone="SOYBEAN CLONE ID: Gm-c1061-5223"
/tissue_type="mature flowers of field grown plants"
/lab_host="DH10B"
/note="Vector: pBluescript II SK+; Site:1: EcoRI; Site:2:
XhoI; The cDNA library was constructed from mRNA isolated
from mature flowers of field grown plants for the cultivar
Raiden. Complementary DNA was synthesized from mRNA using
a primer consisting of a poly(dT) sequence with a XhoI
restriction site. EcoRI adapters were ligated to the
blunt-ended cDNA fragments followed by XhoI digestion. The
cDNA fragments were directionally cloned into the
EcoRI-XhoI restriction site of the pBluescript vector. The
ligated cDNA fragments were transformed into DH10B host
cells (GibcoBRL). This library was constructed in the
laboratory of Dr. Randy Shoemaker."
BASE COUNT      191 a      128 c      111 g      130 t

```



DB	613	CAATATATGGAGCATGTGAGGATA	635
RESULT 14			
LOCUS			
DEFINITION	BE057370	476 bp	mRNA
ACCESSION	BE057370	476 bp	mRNA
VERSION	BE057370.1	GI:8401736	EST
SOURCE	soybean.		
ORGANISM	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.		
REFERENCE	1 (bases 1 to 476)		
AUTHORS	Shoemaker,R., Keim,P., Vodkin,L., Erpeliding,J., Corvelli,V., Khanna,A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers,I., Person,B., Swaller,T., Gibbons,M., Page,D., Harvey,N., Schurk,R., Rittler,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R., Waterston,R. and Wilson,R.		
TITLE	Public Soybean EST Project		
JOURNAL	Unpublished (1999)		
COMMENT	Contact: Shoemaker R/Public soybean EST project Public Soybean EST Project Washington University School of Medicine 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA Tel: 314 286 1800 Fax: 314 286 1810 Email: est@watson.wustl.edu This clone is available through: ResGen, Invitrogen Corp. 2130 South Memorial Parkway Huntsville, AL 35801 For further information call: (800)-533-4363 or contact via email: ccluresgen.com Insert Length: 909 Std Error: 0.00 Seq primer: -40RP from Glbco High quality sequence stop: 420. Location/Qualifiers 1. 476 /organism="Glycine max" /db_xref="taxon:3847" /clone="GENOME SYSTEMS CLONE ID: Gm-c1015-8383" /clone_1id="Gm-c1015" /tissue_type="Mature flowers, field grown plants" /lab_host="Xl10-Gold" /note="vector: pluescript II XR; Site_1: EcoRI; Site_2: XhoI; This cDNA library was constructed from mRNA isolated from mature flowers of field grown plants. The cDNA library was prepared using the Stratagene pluescript II XR cDNA library construction kit. Complementary DNA was synthesized from mRNA using a primer consisting of a poly (dT) sequence with a XhoI restriction site. EcoRI adaptets were ligated to the blunt-ended cDNA fragments followed by XhoI digestion. The cDNA fragments were directionally cloned into the EcoRI-XhoI restriction site of the pluescript vector. The ligated cDNA fragments were transformed into Xl10-Gold host cells. This library was constructed by Dr. Randy Shoemaker and Dr. John Erpeliding."		
BASE COUNT	145 a	113 c	116 g 102 t
ORIGIN			
Query Match	42.1%	Score 363.6;	DB 10; Length 476;
Best Local Similarity	87.8%;	Pred. No. 7.1e-45;	
Matches 409;	Conservative 0;	Mismatches 54;	Indels 3; Gaps 1;
DB	48	GCATACGCTCTCAATCTCTGAAGTGAAGAAAAGGCCCTTGACATGTGCAAGACACTTGA	107
	11	GCATACGCTCTCATATCTCTGAAGTGAAGAAAAGGCCCATGTGACATGTGAAAGAACTTAA	70

QY	108	TCTTGATGAACCTATATTTGCCAATCATGGGGAAAGGTGTTTGGAACTCTTTGGCCAAAGCTG	167
Db	71	TCCTGATCACTATATTTGCCAATCACGGGGAAAGGGTTTGGAACTCTTTGGCCAAAGCTG	130
QY	168	CTGGCTCTCAACAGTATACGGAAGAAGTTCGCCGGCTAAAGTGCTTAAATTAACCTCCGCTC	227
Db	131	CTGAGCTTAAACGTAACCGGAAGAAGTTCCGGCTCCGGTGGCTAAACTACTCCGCTC	190
QY	228	ATGTAGTAAGAGGCGAATATTTACACCCGAGAAACAATTTGATTATGAGGCTCCACGCA	287
Db	191	ATGTAGTAAGAGGCGAATATTTACACCCGAGAAACAATTTGATTATGAGGCTCCATGCA	250
QY	288	AGTGGGGAACAGGAGGATCCAAATTTGGCAGCAATCTACCTGGAAGGACGATATATGA	347
Db	251	AGTGGGGAACAGGAGGATCCAAATTTGGCAGCAATCTACCTGGAAGGACGATATATGA	310
QY	348	TCAAGAACTATTGGAGAGCAAGATCCAGAAAGCAATCAGCAAGCT---GAGAACTTTC	404
Db	311	TTAAGAACTACTGGAGAGCAAGATCCAGAAAGCAATCAGCAAGCTTCCAGAGCTTCC	370
QY	405	AGCAACACAGTATATATTTCTGAGATTAATGATCACCAGCTAGCACTGACCATGTTT	464
Db	371	AGCAACACAGTATATATTTCTGAGATTAATGATCACCAGCTAGCACTGACCAAGTGT	430
QY	465	CCACCATGCTGAGCCCATGAGATGATTTCTCCACCTCGTTATCA	510
Db	431	CCACCATGCTGAGCCCATGAGATGATTTCTCCACCTCGTTATCA	476
RESULT 15			
LOCUS	BM528383		
DEFINITION	BM528383	542 bp	mRNA
	sequence.	54157509.Y1	Gm-cl061 Glycine max cDNA clone SOYBEAN CLONE ID: Gm-cl061-3330 5' similar to TR:Q39028 Q39028 ATMB2. ;, mRNA
ACCESSION	BM528383		
VERSION	BM528383.1	GI:18734780	
KEYWORDS			
SOURCE			
ORGANISM			
			soybean.
			Glycine max
			Eukaryote: Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eustosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.
REFERENCE			
AUTHORS			1 (bases 1 to 542)
			Shoemaker,R., Reim,P., Vodkin,L., Erpelidng,J., Coreyell,V., Khanna
			A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C.,
			Wylie,T., Underwood,K., Stepien,M., Theising,B., Allen,M., Bowers
			'Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk
			'R., Riteier,E., Kohn,S., Shih,T., Jackson,Y., Cardenas,M., McCann
			,R., Waterson,R. and Wilson,R.
TITLE			Public Soybean EST Project
COMMENT			Unpublished (1999)
			Contact: Shoemaker R/Public Soybean EST Project
			Public Soybean EST Project
			Washington University School of Medicine
			4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
			Tel: 314 286 1800
			Fax: 314 286 1810
			Email: est@watson.wustl.edu
			This clone is available through: Resgen, Invitrogen Corp. 2130
			South Memorial Parkway Hunttsville, AL 35801 For further information
			call: (800)-533-4363 or contact: cnu@resgen.com web site:
			www.resgen.com
			Seq primer: -40RP from Gldco
			High quality sequence stop: 423.
FEATURES			
source			
			1..542
			/organism="Glycine max"
			/db_xref="taxon:3847"
			/clone="SOYBEAN CLONE ID: Gm-cl061-3330"
			/clone_11d="Gm-cl061"

/tissue\_type="mature flowers of field grown plants"  
/lab\_host="DH10B"  
/note="Vector: Bluescript II SK+; Site 1: EcoRI; Site 2:  
XhoI; The cDNA library was constructed from mRNA isolated  
from mature flowers of field grown plants for the cultivar  
Raiden. Complementary DNA was synthesized from mRNA using  
a primer consisting of a poly(dT) sequence with a XhoI  
restriction site. EcoRI adapters were ligated to the  
blunt-ended cDNA fragments followed by XhoI digestion. The  
cDNA fragments were directionally cloned into the  
EcoRI-XhoI restriction site of the Bluescript vector. The  
ligated cDNA fragments were transformed into DH10B host  
cells (GibcoBRL). This library was constructed in the  
laboratory of Dr. Randy Shoemaker."

BASE COUNT 191 a 121 c 107 g 123 t  
ORIGIN

Query Match 40.2%; Score 347; DB 13; Length 542;  
Best Local Similarity 78.3%; Pred. No. 1.9e-42;  
Matches 416; Conservative 0; Mismatches 115; Indels 0; Gaps 0;

QY 28 AATGATATAAAACACAGCTGTAAGACGCTCAAGATCCTGAAGTGAAGAAAGGCGCTTG 87  
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||  
Db 5 AATGACAAAAAACACAGCTGTAAGACGCTCAAGATCCTGAAGTGAAGAAAGGCGCATG 64  
QY 88 GACAAATGGAAGAGACTTGATCTGATGAACATATATGCAAAATCATGGGAAGGTGTTG 147  
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||  
Db 65 GACCATGGAAGAGACTTGATCTGATGAACATATATGCAAAATCATGGGAAGGTGTTG 124  
QY 148 GAATCTTTGGCCAAAGCTGCTGCTCAAAAGCTTAACGAAAGAGTTCGCGCTAAGGTG 207  
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||  
Db 125 GAATCTCTTGAAGCAAGCTTCTGCTCAAAAGCTTAACGAAAGAGTTCGCGCTGCTG 184  
QY 208 GCTAAATTTACCTCCGCTCTGATGTTAGAGAGGGAATATTTACACCGAGGAACACTTTT 267  
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||  
Db 185 GCTAAACTACCTTGCTCTGATGTTAGAGAGGGAACATTTACACCGAGGAACACTTTT 244  
QY 268 GATTATGAGAGCTCCAGCAAGTGGGGAACAGTGGTCCAAAATTTGCCAAGCATCTAC 327  
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||  
Db 245 GATCATAGAACTTCATGCAAGTGGGCAATAGTGGTCCAAAATTTGCCAAGCATCTTC 304  
QY 328 TGAAGAGACTGATTAATGATCAAGACTATTTGAGAGACAAAGATCCAGAACATCA 387  
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||  
Db 305 AGGAAGAACTGACAAATGAGATTAGAACTTCTGAGAAACAAAGATCCAAAAGCATTTA 364  
QY 388 GCAAGCTGAGAACTTTGAGCAACAGATAGTAATATTCTGAGATAATGATCACCAGC 447  
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||  
Db 365 GCAAGCTGAGAACTTCACAAACATGTAATTCAGAGAAATATGATCATCAAGCAAGCAC 424  
QY 448 TAGCACTAGCCATGTTTCCACCATGCTGAGCCCATGAGATGATTTCCACCCTGTTA 507  
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||  
Db 425 TAGTACTAGCAAAAGTGTCCACCATGACATCCAAATAGAGACTTCTCTCACCCCTCATA 484  
QY 508 TCAAGGAATGTTAGGCAATTTCAACTCACTGCCCTACAAATTAATCTCGA 558  
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||  
Db 485 CCAAGGAACCTTTGAGCCATTTCAACCTCAATTCCTTACATCATGATCA 535

Search completed: February 18, 2003, 09:19:17  
Job time : 1690 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

## OM protein - nucleic search, using frame\_plus.p2n model

Run on: February 18, 2003, 10:32:35 ; Search time 47 Seconds

(without alignments)  
1344.158 million cell updates/sec

Title: US-10-021-811-36

Perfect score: 1138  
Sequence: 1 MDRKQCKTSQDEVRKGPW.....NNINYSMEDSMOLLNED 206

## Scoring table:

BLOSUM62	
Xgapop 10.0 , Xgapext 0.5	
Ygapop 10.0 , Ygapext 0.5	
Fgapop 6.0 , Fgapext 7.0	
Delop 6.0 , Delext 7.0	

Searched: 441362 seqs, 15338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

## Command line parameters:

-MODEL-frame+.p2n.model -DEV-x1h  
-O=/cgn2\_1/USPTO.spool/US10021811/rnat\_12022003\_090134\_9582/app\_query.fasta\_1.391  
-DB-Issued\_Patents\_NA -OPMT=fastap -SUFFIX=p2n.rni -MINMATCH=0.1 -LOOPCTL=0  
-LOOPEXT=0 -UNITS-bits -START=1 -END=-1 -MATRIX-blosum62 -TRANS-human40.cdi  
-LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTPMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000  
-USER=US10021811\_9CGN\_1.1\_17\_Enunat\_12022003\_090134\_9582 -NCPV=6 -ICPV=3  
-NO\_XLPXY -NO\_MMAP -LARGEQUERY -NDC SCORES=0 -MAIT -LONGLOG -DEV\_TIMEOUT=120  
-WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

## Database :

Issued\_Patents\_NA.\*  
1: /cgn2\_6/prodata/1/lna/5A\_COMB.seq:\*  
2: /cgn2\_6/prodata/1/lna/5B\_COMB.seq:\*  
3: /cgn2\_6/prodata/1/lna/6A\_COMB.seq:\*  
4: /cgn2\_6/prodata/1/lna/6B\_COMB.seq:\*  
5: /cgn2\_6/prodata/1/lna/6C\_COMB.seq:\*  
6: /cgn2\_6/prodata/1/lna/6D\_COMB.seq:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	396	34.8	2220	4	US-08-997-251-1
2	386	33.9	2352	4	US-08-997-251-1
3	382.5	33.6	1344	2	US-08-722-6268-1
4	305.5	26.8	1035	5	PCT-US93-06251-90
5	305.5	26.8	3225	5	US-08-306-691B-45
6	305.5	26.8	3225	5	PCT-US93-06251-91
7	305.5	26.8	3230	5	PCT-US93-06251-89
8	296	26.0	1897	4	US-09-167-322-10
9	293	25.7	4880	4	US-09-402-925-5
10	291.5	25.6	6702	4	US-09-402-925-1
11	291.5	25.6	6775	4	US-09-402-929-1
12	271.5	23.9	2638	1	US-08-306-691B-46

13	248	21.8	4059	2	US-08-485-139-1	Sequence 1, Appl
14	248	21.8	4059	3	US-08-750-357-1	Sequence 1, Appl
15	248	21.8	4824	2	US-08-485-139-5	Sequence 5, Appl
16	248	21.8	4824	3	US-08-750-357-5	Sequence 5, Appl
17	185.5	16.3	469	3	US-09-008-979A-2	Sequence 2, Appl
18	185.5	16.3	469	4	US-09-460-618-2	Sequence 2, Appl
19	185.5	16.3	469	4	US-09-310-235B-2	Sequence 2, Appl
20	165.5	14.5	5889	4	US-09-402-929-3	Sequence 3, Appl
21	125	11.0	785	3	US-09-008-979A-6	Sequence 6, Appl
22	125	11.0	785	4	US-09-460-618-6	Sequence 6, Appl
23	125	11.0	785	4	US-09-310-235B-6	Sequence 6, Appl
24	119	10.5	2837	4	US-09-156-316-11	Sequence 11, Appl
25	105.5	9.3	1116	4	US-08-928-941D-17	Sequence 17, Appl
26	105.5	9.3	1116	4	US-09-280-590A-17	Sequence 17, Appl
27	105.5	9.3	2903	4	US-08-928-941D-2	Sequence 2, Appl
28	105.5	9.3	2903	4	US-08-928-941D-3	Sequence 2, Appl
29	105.5	9.3	2903	4	US-09-280-590A-2	Sequence 2, Appl
30	105.5	9.3	2903	4	US-09-280-590A-3	Sequence 3, Appl
31	105.5	9.3	3767	4	US-08-928-941D-28	Sequence 28, Appl
32	105.5	9.3	3767	4	US-08-928-941D-30	Sequence 30, Appl
33	105.5	9.3	3767	4	US-09-280-590A-28	Sequence 28, Appl
34	105.5	9.3	3767	4	US-09-280-590A-30	Sequence 30, Appl
35	96.5	8.5	282	2	US-08-814-030-2	Sequence 2, Appl
36	93	8.2	336	1	US-08-248-474-105	Sequence 105, App
37	93	8.2	336	3	US-08-756-849-105	Sequence 105, App
38	84.5	7.4	3834	4	US-09-221-017B-663	Sequence 663, App
39	83.5	7.3	2581	4	US-09-370-838-66	Sequence 66, Appl
40	82.5	7.2	3642	3	US-08-946-026-16	Sequence 16, Appl
41	81.5	7.2	3627	1	US-08-447-500-23	Sequence 23, Appl
42	81.5	7.2	3627	1	US-08-454-097-23	Sequence 23, Appl
43	81.5	7.2	3627	3	US-08-453-866-23	Sequence 23, Appl
44	81.5	7.2	3627	3	US-08-185-359-23	Sequence 23, Appl
45	81.5	7.2	3628	1	US-08-468-036-17	Sequence 17, Appl

## ALIGNMENTS

RESULT 1  
US-08-997-251-1  
; Sequence 1, Application US/08997251  
; Patent No. 6271440  
GENERAL INFORMATION:  
; APPLICANT: GUBLER, FRANZ J.  
; APPLICANT: JACOBSEN, JOHN V.  
; TITLE OF INVENTION: PLANT REGULATORY PROTEINS III  
; NUMBER OF SEQUENCES: 17  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Greenlee, Winner and Sullivan, P.C.  
; STREET: 5370 Manhattan Boulevard  
; CITY: Boulder  
; STATE: CO  
; COUNTRY: US  
; ZIP: 80303  
COMPUTER READABLE FORM:  
; MEDIUM TYPE: floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/997,251  
; FILING DATE: 23-DEC-1997  
CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: WO AU96/00383  
; FILING DATE: 21-JUN-1996  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: AU PN6470/95  
; FILING DATE: 09-NOV-1995  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: AU PN3779/95  
; FILING DATE: 23-JUN-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Winner, Ellen P.

```

: REGISTRATION NUMBER: 28,547
: REFERENCE/DOCKET NUMBER: 110-97
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (303) 499-8080
: TELEFAX: (303) 499-8089
: INFORMATION FOR SEQ ID NO: 1:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 2220 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: cDNA
: ORIGINAL SOURCE:
: ORGANISM: Hordeum vulgare
: STRAIN: Himalaya
: DEVELOPMENTAL STAGE: Seed
: TISSUE TYPE: Aleurone layer
: IMMEDIATE SOURCE:
: CLONE: HVCAMyb
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 275..1933
: US-08-997-251-1

Alignment Scores:
Pred. No.: 1,47e-38 Length: 2220
Score: 396.00 Matches: 70
Percent Similarity: 69.50% Conservative: 28
Best Local Similarity: 49.65% Mismatches: 31
Query Match: 34.80% Indels: 12
DB: 4 Gaps: 1

US-10-021-811-36 (1-206) x US-08-997-251-1 (1-2220)
QY 13 ProgluValArgIysGlyProTTrpThrMetGluGluAspLeuIleuMetAsnTyrIle 32
DB 392 CCGCCTGTGAAGAGGGCCCTGGACCTCCGCGAGGAGCGCCATCTCGTGCACGTG 451
QY 33 AlaAsnHisGlyGluGlyValTTrpAsnSerLeuAlaIysAlaIysLeuLysArgAsn 52
DB 452 AAGAAAGCAGCGGAGGAGGAACTGGAAACCGGTGCAGAAAGAACACCGGCTTCCGGTGC 511
QY 53 GlyLysSerCysArgLeuArgTTrpLeuAsnTyrLeuArgProAspValArgArgIysAsn 72
DB 512 GGCAAGAGCTGCCGCTCCGCTGGGCGAACCCTCAGGCCCAACCTCAAGAGGGGCC 571
QY 73 IleThrProGluGluGlnLeuLeuIleMetGluLeuHisAlaLysTTrpGlyAsnArgTTrp 92
DB 572 TTCACCCCGAGAGGAGAGAGCTCATCTCCAGCTCCACATCCAGATGGCAACAAGTGG 631
QY 93 SerLysIleAlaLysHisLeuProGlyArgTTrpAspAsnGluIleLysAsnTyrTTrpArg 112
DB 632 GCTCGAATGCCCGCTCATTTGCCAGGGCGTACTGATTAATAAGAAATTAAGTACTGGAAC 691
QY 113 ThrArgIleGlnLys-----HisIleLys 120
DB 652 ACTCGAATAAAGATGTGCAGCGAGCCGTTGCCAATATATCTCGTAGTGTATGCAT 751
QY 121 GlnAlaGluAspHieGlnGlnInSerSerAsnAsnSerGluIleAsnAspHisGlnAla 140
DB 752 CAATCTCAAAATGAAGATCAGCAGAGGGCTCCAGCATTTCAACTGGCGGAGAACTTTCC 811
QY 141 Ser 141
DB 812 AGT 814

RESULT 2
US-08-997-251-3
: Sequence 3, Application US/08997251
: Patent No. 6271440
: GENERAL INFORMATION:
: APPLICANT: GUBLER, FRANZ J.
: APPLICANT: JACOBSEN, JOHN V.

```

```

: TITLE OF INVENTION: PLANT REGULATORY PROTEINS III
: NUMBER OF SEQUENCES: 17
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Greenlee, Winner and Sullivan, P.C.
: STREET: 5370 Manhattan Boulevard
: CITY: Boulder
: STATE: CO
: COUNTRY: US
: ZIP: 80303
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patent Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/997,251
: FILING DATE: 23-DEC-1997
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: WO A096/00383
: FILING DATE: 21-JUN-1996
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: AU PN6470/95
: FILING DATE: 09-NOV-1995
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: AU PN3779/95
: FILING DATE: 23-JUN-1995
: ATTORNEY/AGENT INFORMATION:
: NAME: Winner, Ellen P.
: REGISTRATION NUMBER: 28,547
: REFERENCE/DOCKET NUMBER: 110-97
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (303) 499-8080
: TELEFAX: (303) 499-8089
: INFORMATION FOR SEQ ID NO: 3:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 2352 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: cDNA
: ORIGINAL SOURCE:
: ORGANISM: Oryza sativa
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 396..2054
: US-08-997-251-3

Alignment Scores:
Pred. No.: 2,69e-37 Length: 2352
Score: 386.00 Matches: 68
Percent Similarity: 69.57% Conservative: 28
Best Local Similarity: 49.28% Mismatches: 30
Query Match: 33.92% Indels: 12
DB: 4 Gaps: 1

US-10-021-811-36 (1-206) x US-08-997-251-3 (1-2352)
QY 13 ProgluValArgIysGlyProTTrpThrMetGluGluAspLeuIleuMetAsnTyrIle 32
DB 507 CCGCCTGTGAAGAGGGCCCATGAGCTGCGCGAGGAGCGCCATCTCGTGCACATCAGTGG 566
QY 33 AlaAsnHisGlyGluGlyValTTrpAsnSerLeuAlaIysAlaIysLeuLysArgAsn 52
DB 567 AAGAAAGCAGCGGAGGAGGAACTGGAAACCGGTGCAGAAAGAACACCGGCTTCCGGTGC 626
QY 53 GlyLysSerCysArgLeuArgTTrpLeuAsnTyrLeuArgProAspValArgArgIysAsn 72
DB 627 GGCAAGAGCTGCCGCTCCGCTGGGCGAACCCTGAGGCCCAAGTCAAGAGGGGCC 686
QY 73 IleThrProGluGluGlnLeuLeuIleMetGluLeuHisAlaLysTTrpGlyAsnArgTTrp 92
DB 687 TTCACCCCGAGAGGAGAGAGCTCATCTCCAGCTCCACATGCGGAGAACTGGTGG 746

```

QY 93 SerlystlealalyshisleuproglyarqthraspasngluilelysasnTYTTPArq 112  
Db 747 GCTGGAGATGCGCTCATTTTCAGGGCGACGTATATGAATAAAGAAATTACTGGAAT 806  
QY 113 ThrqTgllleglnlys-----Hisllelys 120  
Db 807 ACTGGAATAAAGACATGCCAGCGAGCTGGCCTACCCATCTATCTCTACCAGCGTATGCAAT 866  
QY 121 GlnaIaGuaenphegInglInserSerAsnAsnsergluileAsnAspHis 138  
Db 867 CAATCCTCAATGAAGATCAGCACTGCTCCAGTATTTTGACCTGGCGAGAAAT 920

RESULT 3  
US-08-722-626B-1  
Sequence 1, Application US/08722626B  
Patent No. 5939601  
GENERAL INFORMATION:  
APPLICANT: Yang, Yilong  
APPLICANT: Klessig, Daniel, F.  
TITLE OF INVENTION: NOVEL GENES ASSOCIATED WITH ENHANCED  
DISEASE RESISTANCE IN PLANTS  
NUMBER OF SEQUENCES: 9  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Dann, Dorfman, Herrell and Skillman  
STREET: 1601 Market Street Suite 720  
CITY: Philadelphia  
STATE: PA  
COUNTRY: USA  
ZIP: 19103-2307  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq Version 1.5  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/722,626B  
FILING DATE: 27-SEP-1996  
CLASSIFICATION: 800  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Pat Hagan  
REGISTRATION NUMBER: 27,643  
REFERENCE/DOCKET NUMBER: 97-0010  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 215 563-4100  
TELEFAX: 215 563-4044  
TELEX:  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1344 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
FRAGMENT TYPE:  
ORIGINAL SOURCE:  
FEATURE:  
NAME/KEY: Coding Sequence  
LOCATION: 148...981  
OTHER INFORMATION:  
US-08-722-626B-1

Alignment Scores:  
Pred. No.: 3,09e-37 Length: 1344  
Score: 382.50 Matches: 88  
Percent Similarity: 54.38% Conservative: 30  
Best Local Similarity: 40.55% Mismatches: 60  
Query Match: 33.61% Indels: 39

DB: 2 Gaps: 6  
US-10-021-811-36 (1-206) x US-08-722-626B-1 (1-1344)  
QY 15 ValArgLysglYProtrpThrMetgluAspLeuileuMetAsnTYTllealaasn 34  
Db 181 CTGAAAAAAGGCCCATGATTCCTGAAGAGATCAGATTCCTCATCTTCATTCAAACT 240  
QY 35 HisglYgluIlyValTTPAsnSerleuAlaIyAlaIglYleuysrAsnnglylys 54  
Db 241 AATGGCGATGGCAACAGCGGAGCCCTCCCAAAAGCGTGCATATGGATGCGGGAG 300  
QY 55 SerCysArgLeuArgTrpLeuAsnTYrLeuArgProAspValArgArgGlyAsnIleThr 74  
Db 301 AGTTGGAGACTGGCTGGAGCATTTATTTGGCCAGATATTAAGAGGGGAATTTACAC 360  
QY 75 ProgluIgluIleuIleuIleuIleuIleuIleuHisAlaIyTrpIlysnArgTrpSerIys 94  
Db 361 AAGGAAGAACAGCAAAATATTCAGATTACATGAATGCTGGCAATGATGCTGCA 420  
QY 95 IleAlaIyshIsleuproglyarqthraspasngluilelysasnTYTTPArqThrarq 114  
Db 421 ATAGCAGCAAAATTTACCGAGCAGACAGCAAGAAATGAAATGTTTGGCACACCCAC 480  
QY 115 IleGlnIyshIslelys-----GlnAla 122  
Db 481 TTGAAGAGAGAGCTCAAGATTATAAGCTCTCAGAACTCCAAAAGACATCCAAATCC 540  
QY 123 GluAsnPhedIn-----GlnIleSerSerAsnAsnsergluile--- 135  
Db 541 AACAAATCATGATTCCAAGGCTCTACTACTCTGAAATCATCCAAATATCTGATCTACT 600  
QY 136 -----AsnAspHisGlnAla 140  
Db 601 ATTATTAAATACAAAAACAAATGATAGCCAGCTAGTCTCCTACCTACACCCAAAT 660  
QY 141 SerThSer---HisValSerThrMetAlaGluProMetGluMetIySerProProCys 159  
Db 661 TCATCTAGTACTGAATGTCACACTGACACTAGTCAGATGATCAATAGTGTGAT 720  
QY 160 TYrGlnIyMetLeuGluProPheSerThrGlnPheProThrIleAsnProAspGlnSer 179  
Db 721 AACGAAGAAGTAAATGAGATCG---TCCGAGTATTTTCCAGACATGAT-----GAG 768  
QY 180 SerCysCysThrasnAsnAsnAsnIleAsnTYrTrpSerMetGluAsp 196  
Db 769 AGTTTGGAGCGAGCATTAACACGACAAATTAAGTGAAGTGAAGTGAAGTGAAGT 819

RESULT 4  
PCT-US93-06251-90  
Sequence 90, Application PC/TUS9306251  
GENERAL INFORMATION:  
APPLICANT: Wickstrom, Eric and Rife, Jason P.  
TITLE OF INVENTION: Trivalent Synthesis of Oligonucleotides Containing  
Stereospecific Alkylphosphonates and Arylphosphonates  
NUMBER OF SEQUENCES: 93  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: SCUDLY, SCOTT, MURPHY & PRESSER  
STREET: 400 Garden City Plaza  
CITY: Garden City  
STATE: NY  
COUNTRY: USA  
ZIP: 11530  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US93/06251  
FILING DATE: 19930630  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:

NAME: Digilio, Frank S.  
 REGISTRATION NUMBER: 31,346  
 REFERENCE/DOCKET NUMBER: 8586  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 516-742-4343  
 TELEFAX: 516-742-4366  
 TELEX: 230 901 SANS UR  
 INFORMATION FOR SEQ ID NO: 90:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 1035 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: double  
 TOPOLOGY: linear  
 MOLECULE TYPE: DNA (genomic)  
 PCT-0593-06251-90

Alignment Scores:  
 Pred. No.: 5,58e-28 Length: 1035  
 Score: 305.50 Matches: 70  
 Percent Similarity: 59.34% Conservative: 38  
 Best Local Similarity: 38.46% Mismatches: 59  
 Query Match: 26.85% Indels: 15  
 Gaps: 7  
 DB:

US-10-021-811-36 (1-206) x PCT-US93-06251-90 (1-1035)

OY 12 ASPPROGLUVALARGLYSGLYPROTPRTHMETGLUGLUASPLEUULEUUMETASNTYR 31  
 DB 121 AACCTGAGCTCATACAAAGGCTCTTGACCAAGAGAGATCAGAGGTATAGACTT 180  
 OY 32 ILEALAASNHISGLYGLUGLYVALTRPASNSEUUALALYSALALAGLYLEULYS--- 50  
 DB 181 GTACGAAATACGGTCCCAACGTTGCTGTATTGTGCCAAG-----CACCTTAAAGGG 234  
 OY 51 ARGASNGLYLVSERCYSARGLEUARGTPLEUASNTYRLEUARGPROASPVALARARG 70  
 DB 235 AGAATTGGAAAAAATGAGGAGAGAGGTGACATACCACTTGAAATCCAGAGTTAAGAAA 294  
 OY 71 GLYASNILETHPRROGLUGLUGLULEUULEULEUULEUULEUULEUULEUULEUULEU 90  
 DB 295 ACCTCTGTGACAGAGAGAGAGAGACGAATTTATTCACGACACACAGAGACTGGGAAAC 354  
 OY 91 ARGTPSERLYSILEALALYSIHISLEUPROGLYARGTPHASPANGLULILEYASNTYR 110  
 DB 355 AGATGGGAGAAATGCAAGACTACTGCTGACGAACTGATATGCTATCAGAACACAC 414  
 OY 111 TRPARGTTHARGILEGLINLISHISILEYSGINLALAGLUANPHEGLINGLISER 130  
 DB 415 TGGAAATTCACATGCTGCGAGAGTCCACACAG--GAAGTTATCTGACGAGCTTCA 471  
 OY 131 ASNASNSERGLU-----ILEASNAIPHISGLINALASERTHRSERHISVALSERTHMET 148  
 DB 472 AAAGCCACCCAGCCAGTGGCCACAAGCTTCCAGAAAGACAGCATTTGATGGGTTT 531  
 OY 149 AAGIUPROMETGLUMETYSERPROPCYSTYRGLINGLIMETLEUGLUPROPHESER 168  
 DB 532 GCTCGAG-----GCTCGCCCTACAGCTCA-----CTCCCTGCCACT 567  
 OY 169 THRGINPHEPROTHILEASNPPOASPLINSERCYSCTYRTHASNPASNPASNPASN 188  
 DB 568 GGCCAG---CCCACTGTTAACAACAGACTATTCATTACACATTTCTGAAGCACAATAAT 624  
 OY 189 ILEASN 190  
 DB 625 GTCTCC 630

RESULT 5  
 US-08-306-691B-45  
 Sequence 45, Application US/08306691B  
 Patent No. 5734039  
 GENERAL INFORMATION:  
 APPLICANT: Calabretta, Bruno  
 APPLICANT: Skorski, Tomasz

TITLE OF INVENTION: ANTISENSE  
 NUMBER OF INVENTION: OLIGONUCLEOTIDES TARGETING COOPERATING ONCOGENES  
 NUMBER OF SEQUENCES: 55  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Seidel, Gonda, Lavoragna & Monaco, P. C.  
 STREET: Two Penn Center, Suite 1800  
 CITY: Philadelphia  
 STATE: Pennsylvania  
 COUNTRY: U.S.A.  
 ZIP: 19102  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Diskette, 3.50 inch, 720 KB  
 COMPUTER: IBM PS/2  
 OPERATING SYSTEM: MS-DOS  
 SOFTWARE: WordPerfect 5.1  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/306,691B  
 FILING DATE: September 15, 1994  
 CLASSIFICATION: 514  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER:  
 FILING DATE:  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Monaco, Daniel A.  
 REGISTRATION NUMBER: 30,480  
 REFERENCE/DOCKET NUMBER: 8321-8  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (215) 568-8383  
 TELEFAX: (215) 568-5549  
 TELEX: NO. 5734039e  
 INFORMATION FOR SEQ ID NO: 45:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 3225 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: double  
 TOPOLOGY: linear  
 US-08-306-691B-45

Alignment Scores:  
 Pred. No.: 3.13e-27 Length: 3225  
 Score: 305.50 Matches: 70  
 Percent Similarity: 59.34% Conservative: 38  
 Best Local Similarity: 38.46% Mismatches: 59  
 Query Match: 26.85% Indels: 15  
 Gaps: 7  
 DB:

US-10-021-811-36 (1-206) x US-08-306-691B-45 (1-3225)

OY 12 ASPPROGLUVALARGLYSGLYPROTPRTHMETGLUGLUASPLEUULEUUMETASNTYR 31  
 DB 372 AACCTGAGCTCATACAAAGGCTCTTGACCAAGAGAGATCAGAGGTATAGACTT 431  
 OY 32 ILEALAASNHISGLYGLUGLYVALTRPASNSEUUALALYSALALAGLYLEULYS--- 50  
 DB 432 GTACGAAATACGGTCCCAACGTTGCTGTATTGTGCCAAG-----CACCTTAAAGGG 485  
 OY 51 ARGASNGLYLVSERCYSARGLEUARGTPLEUASNTYRLEUARGPROASPVALARARG 70  
 DB 486 AGAATTGGAAAAAATGAGGAGAGGTGACATACCACTTGAAATCCAGAGTTAAGAAA 545  
 OY 71 GLYASNILETHPRROGLUGLUGLULEUULEULEUULEUULEUULEUULEUULEUULEU 90  
 DB 546 ACCTCTGTGACAGAGAGAGACGAATTTATTCACGACACACAGAGACTGGGAAAC 605  
 OY 91 ARGTPSERLYSILEALALYSIHISLEUPROGLYARGTPHASPANGLULILEYASNTYR 110  
 DB 606 AGATGGGAGAAATGCAAGACTACTGCTGACGAGATGATATGCTTCAAGAACCCAC 665  
 OY 111 TRPARGTTHARGILEGLINLISHISILEYSGINLALAGLUANPHEGLINGLISER 130  
 DB 666 TGGAAATTCACATGCTGCGAGAGTCCACACAG--GAAGTTATCTGACGAGCTTCA 722  
 OY 131 ASNASNSERGLU-----ILEASNAIPHISGLINALASERTHRSERHISVALSERTHMET 148



Db 723 AAAGCCAGCCAGCCAGCAGTGGCCACCAAGCTTCACGAAAGAACTCATTTGATGGGTTT 782  
QY 149 AAlaGluProMetGluMetTySerProCysTyGlnGlyMetLeuGluProPheSer 168  
Db 783 GCTCAG-----GCTCGGCTACAGCTCAA-----CTCCCTGCCACT 818  
QY 169 ThrGlnPheProThrIleAsnProAspGlnSerSerCysCysThrAsnAspAsnAsn 188  
Db 819 GGCCAG---CCCACTGTAAACAAGCACTATTCTATTACCAATTTCTGAAGCACAAAAT 875  
QY 189 Ileasn 190  
Db 876 GTCTCC 881  
RESULT 6  
PCT-US93-06251-91  
Sequence 91, Application PC/TUS9306251  
GENERAL INFORMATION:  
APPLICANT: Wickstrom, Eric and Rife, Jason P.  
TITLE OF INVENTION: Trivalent Synthesis of Oligonucleotides Containing  
TITLE OF INVENTION: Stereospecific Alkylphosphonates and Arylphosphonates  
NUMBER OF SEQUENCES: 93  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: SCULLY, SCOTT, MURPHY & PRESSER  
STREET: 400 Garden City Plaza  
CITY: Garden City  
STATE: NY  
COUNTRY: USA  
ZIP: 11530  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US93/06251  
FILING DATE: 19930630  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Digiglio, Frank S.  
REGISTRATION NUMBER: 31,346  
REFERENCE/DOCKET NUMBER: 8586  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 516-742-4343  
TELEFAX: 516-742-4366  
TELEX: 230 901 SANS UR  
INFORMATION FOR SEQ ID NO: 91:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 3225 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
PCT-US93-06251-91  
Alignment Scores:  
Pred. No.: 3,13e-27 Length: 3225  
Score: 305.50 Matches: 70  
Percent Similarity: 59.34% Conservative: 38  
Best Local Similarity: 38.46% Mismatches: 59  
Query Match: 26.85% Indels: 15  
Gaps: 7  
US-10-021-811-36 (1-206) x PCT-US93-06251-91 (1-3225)  
QY 12 AspProGluValAlaGlySerGlyProTPrThrMetGlnGluAspLeuIleuMetAsnTy 31  
Db 372 AACCTGAGCTCAATCAAGGCTCTTGACCAAGAAAGATCAGAGATGATGAGCTT 431  
QY 32 TleAlaAsnHisGlnGlyGluGlyValTyrPasnSerLeuAlaIleAlaGlyLeuLys--- 50  
Db 432 GTACAGAAATAAGCGTCGAAAGCTGTGTCTATTATGCCAAG-----CACTTAAAGGG 485

QY 51 ArgAsnGlyLysSerCysArgLeuArgTyrPleuAsnTyrIleuArgProAspValArgArg 70  
Db 486 AGATTGTGAAATAAGATGTAGGAGAGAGGTGCATTAACCACTTGATTCACAACTTAAGAA 545  
QY 71 GlyAsnIleThrProGluGluGlnLeuIleMetGlnLeuHisAlaLysTrrGlyAsn 90  
Db 546 ACCTCGTGACAMAGACAGACAGATTAATTATTACAGCAGCACAGACAGATGGGAAAC 605  
QY 91 ArgTyrSerLysIleAlaLysHisLeuProGlyArgTyrAspAsnGluIleLysAsnTy 110  
Db 606 AGATGGGCAAAATCCCAAGCTACTGCTGGACAGACTGATTAATGCTATCAAGAACAC 665  
QY 111 TrrArgThrArgIleGlnLysHisIleLysGlnAlaGluAsnPheGlnGlnSerSer 130  
Db 666 TGGAATTTCAATATGCGTCGGAAGGTGCAACAG---GAAGTATCTGAGAGAGCTTCA 722  
QY 131 AsnAsnSerGlu-----IleAsnAspHisGlnAlaSerThrSerHisValSerThre 148  
Db 723 AAAGCCAGCCAGCCAGCAGTGGCCACACAGCTTCAGAAAGCAAGCATTTGATGGGTTT 782  
QY 149 AAlaGluProMetGluMetTySerProCysTyGlnGlyMetLeuGluProPheSer 168  
Db 783 GCTCAG-----GCTCGGCTACAGCTCAA-----CTCCCTGCCACT 818  
QY 169 ThrGlnPheProThrIleAsnProAspGlnSerSerCysCysThrAsnAspAsnAsn 188  
Db 819 GGCCAG---CCCACTGTAAACAAGCACTATTCTATTACCAATTTCTGAAGCACAAAAT 875  
QY 189 Ileasn 190  
Db 876 GTCTCC 881  
RESULT 7  
PCT-US93-06251-89  
Sequence 89, Application PC/TUS9306251  
GENERAL INFORMATION:  
APPLICANT: Wickstrom, Eric and Rife, Jason P.  
TITLE OF INVENTION: Trivalent Synthesis of Oligonucleotides Containing  
TITLE OF INVENTION: Stereospecific Alkylphosphonates and Arylphosphonates  
NUMBER OF SEQUENCES: 93  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: SCULLY, SCOTT, MURPHY & PRESSER  
STREET: 400 Garden City Plaza  
CITY: Garden City  
STATE: NY  
COUNTRY: USA  
ZIP: 11530  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US93/06251  
FILING DATE: 19930630  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Digiglio, Frank S.  
REGISTRATION NUMBER: 31,346  
REFERENCE/DOCKET NUMBER: 8586  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 516-742-4343  
TELEFAX: 516-742-4366  
TELEX: 230 901 SANS UR  
INFORMATION FOR SEQ ID NO: 89:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 3230 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
PCT-US93-06251-89

Alignment Scores: 3.14e-27 Length: 3230  
Pred. No.: 305.50 Matches: 70  
Score: 59.34% Conservative: 38  
Percent Similarity: 59.34% Mismatches: 59  
Best Local Similarity: 38.46% Indels: 15  
Query Match: 26.85% Gaps: 7

US-10-021-811-36 (1-206) x PCT-US93-06251-89 (1-3230)

QY 12 AspProgluValArgylsGlyProTrpThrMetGluGluAspLeuIleuMetAsnTyr 31  
Db 262 AACCTGAGCTCATCAAGGCTCTGGACCAAGAGATCATAGAGATGAGACTT 321  
QY 32 IleAlaAsnHisGlyGluGlyValTrpAsnSerIleuAlaLysAlaGlyLeuLys-- 50  
Db 322 GTACGAAATACGTCACCAAAAGCTTGCTCTTATGGCAAG-----CACTTAAAGGG 375  
QY 51 ArgAsnGlyLysSerCysArgLeuArgTrpLeuAsnTyrLeuArgProAspValArgArg 70  
Db 376 AGAATTGGAAAAACATGTAGGAGAGAGTGCATACCACTTGAATCCAGAAATTAGAAA 435  
QY 71 GlyAsnIleTrpProGluGluGlnLeuIleuMetGluLeuHisAlaLysTrpGlyAsn 90  
Db 436 AACCTCTGGACAGAGAGAGAGATTTATTCAGGACACAGAGAGCTGGGAGAC 495  
QY 91 ArgTrpSerLysIleAlaLysHisLeuProGlyArgTrpAsnGluIleLysAsnTyr 110  
Db 496 AGATGGGAGAAATCGCAAACTGCTGCTGAGAACTGATATGCTATCAGAAACCCAC 555  
QY 111 TrpArgTrpArgIleGlnLysHisIleLysGlnAlaGluAsnPhoGlnGlnGlnSerSer 130  
Db 556 TGGATTCTACAAATCGTCGAGAGGTGACACAG--GAAGTTATCTGACAGACTCTTCA 612  
QY 131 AsnAsnSerGlu-----IleAsnAspHisGlnAlaSerTrpSerHisValSerThmet 148  
Db 613 AAAGCCAGCCAGCAGCAGTGGCCACAAAGCTTCAGAAAGAACAGTATTGATGGGTTT 672  
QY 149 AlaGluProMetGluMetLysSerProProCysTyrGlnGlyMetLeuGluProPhoSer 168  
Db 673 GCTCAG-----GCTCGGCTTACAGTCA-----CTCCCTGCGCACT 708  
QY 169 ThrGlnPheProThrIleAsnProAspGlnSerSerCysTrpAsnAspAsnAsn 188  
Db 709 GCCCAG---CCACTGTTAACACACACATTCATTATACACATTTCTGAAGCACAAAT 765  
QY 189 IleAsn 190  
Db 766 GTCCTCC 771

RESULT 8  
US-09-167-322-10  
Sequence 10, Application US/09167322  
Patent No. 6365151  
GENERAL INFORMATION:  
APPLICANT: Allegheny University of the Health  
Sciences, Halpern, Michael S.  
England, James M.  
TITLE OF INVENTION: CANCER VACCINE  
NUMBER OF SEQUENCES: 14  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Seidel, Gonda, Lavorgna & Monaco, P.C.  
STREET: Suite 1800, Two Penn Center Plaza  
CITY: Philadelphia  
STATE: PA  
COUNTRY: USA  
ZIP: 19102  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/167, 322  
FILING DATE: 07-Oct-1998  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/US97/00582  
FILING DATE: <Unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: Monaco, Daniel A.  
REGISTRATION NUMBER: 30,480  
REFERENCE/DOCKET NUMBER: 7933-33 PC  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (215) 568-8383  
TELEFAX: (215) 568-5549  
INFORMATION FOR SEQ ID NO: 10:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1897 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
SEQUENCE DESCRIPTION: SEQ ID NO: 10:  
US-09-167-322-10

Alignment Scores: 2.04e-26 Length: 1897  
Pred. No.: 296.00 Matches: 73  
Score: 53.47% Conservative: 35  
Percent Similarity: 53.47% Mismatches: 55  
Best Local Similarity: 36.14% Indels: 40  
Query Match: 26.01% Gaps: 9

US-10-021-811-36 (1-206) x US-09-167-322-10 (1-1897)

QY 12 AspProgluValArgylsGlyProTrpThrMetGluGluAspLeuIleuMetAsnTyr 31  
Db 233 AACCTGAGCTCATCAAGGCTCTGGACCAAGAGATCAAGGATATAGAACAC 292  
QY 32 IleAlaAsnHisGlyGluGlyValTrpAsnSerIleuAlaLysAlaGlyLeuLys-- 50  
Db 293 GTCCGAAATACCGTCCAAAGCGCTGTCGACATGTCTAAG-----CATTTGAAGGA 346  
QY 51 ArgAsnGlyLysSerCysArgLeuArgTrpLeuAsnTyrLeuArgProAspValArgArg 70  
Db 347 AGATTGGAAACAGTGCAGAGAGAGTGCACACACATCTGATATCAGAAAGTGAAGA 406  
QY 71 GlyAsnIleTrpProGluGluGlnLeuIleuMetGluLeuHisAlaLysTrpGlyAsn 90  
Db 407 AACCTCTGGACAGAGAGAGATTTATTTACAGGACACAGAGACTGGGAAAC 466  
QY 91 ArgTrpSerLysIleAlaLysHisLeuProGlyArgTrpAsnGluIleLysAsnTyr 110  
Db 467 AGATGGGAGAAATTCGAAAGTGTGCTGCGACGAGATGATACGCTGTCAAGAACAC 526  
QY 111 TrpArgTrpArgIleGlnLysHisIleLysGlnAlaGluAsnPhoGlnGlnGlnSerSer 130  
Db 527 TGGAAATTCACATGCGCGGAGAGTGCAGAG--GAGGTTATCCCGCAGAGTCTCC 583  
QY 131 Asn-----AsnSerGluIleAsnAspHisGlnAlaSerTrpSerHisValSerThr 147  
Db 584 AAAGCCGCGCGCGCTCGCAACACCGGCTTCACAAAGAC---AGCATCTGATGGC 640  
QY 148 MetAlaGluProMetGluMetLysSerProProCys----- 159  
Db 641 TTGGCC-----CACAAACCACTCGAGGCGCTCCCGGCGCGCCAG 685  
QY 160 -----TyrGlnGlyMetLeuGluProPhoSerThr----- 169  
Db 686 GCCCTCTGGGACAGTACCTACCTTACACATGTGAGGCAAAATGCTCCTGCT 745  
QY 170 GlnPhePro-----ThrIleAsnProAspGlnSerSerCys 181  
Db 746 CAGATCCCATATCCAGTACGATGATTAATATATATATCAAGTCTCCTCAGCC-AGCTCC 804

QY 182 CysThr 183  
DB 805 TGCAC 810

RESULT 9  
US-09-402-929-5  
Sequence 5, Application US/09402929  
Patent No. 6410825

GENERAL INFORMATION:  
APPLICANT: Temple University - Of The Commonwealth System of Higher Education  
APPLICANT: Toscani, Antonio  
APPLICANT: Hatton, Kimi  
APPLICANT: Reddy, E. P.  
TITLE OF INVENTION: A-myb NULL MUTANT TRANSGENIC ANIMALS AND  
TITLE OF INVENTION: USES THEREOF  
NUMBER OF SEQUENCES: 7  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: SEIDEL, GONDA, LAVORGNA & MONACO, P.C.  
STREET: Suite 1800 Two Penn Center Plaza  
CITY: Philadelphia  
STATE: PA  
COUNTRY: U.S.A.  
ZIP: 19102

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/402,929  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/US98/06896  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Monaco, Daniel A.  
REGISTRATION NUMBER: 30,480  
REFERENCE/DOCKET NUMBER: 6056-214 PC  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (215) 568-8383  
TELEFAX: (215) 568-5549  
INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 4880 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-09-402-929-5

Alignment Scores:  
Pred. No.: 2e-25 Length: 4880  
Score: 293.00 Matches: 70  
Percent Similarity: 56.22% Conservative: 34  
Best Local Similarity: 37.84% Mismatches: 57  
Query Match: 25.75% Indels: 24  
Gaps: 8

US-10-021-811-36 (1-206) x US-09-402-929-5 (1-4880)

QY 12 AspprogluValarglysglyProtrpThmerglugluAspleuileuMetAsnTyr 31  
DB 348 AATCTGAAATTTGATTAAGGGCTCTTGACATAAGAGACATCAGCGTTATTGAATTA 407

QY 32 lLeAlaasnhtslgylugluValTTPAsnSerleuAlaAlaaglyLeuLys--- 50  
DB 408 GTTCAGAAATATGGCCAAAAGATGCTTTTATTCAGAAA-----CATTTAAAAAGA 461

QY 51 ArgasnlylysserCysargyleuArgTTPleuAsnTyrleuArgProaspValargarg 70  
DB 462 AGAATATAGCAAGCAGTGTAGAGAAAGATGGCATATCATCTGAATCTCGAGTAAGAAA 521

QY 71 GlyAsnleThrProglugluInleuLeuileuMetgluLeuHisAlaLysTTPglyAsn 90

DB 522 TCTTCCTGGACAGAAAGAGAGACAGCATCATCTATGAAAGACATAGCGGTTGGAAAT 581

QY 91 ArgTTPserlyslleAlaLysHsleuProgluYArgThrAspAsnGluileLysAsnTyr 110  
DB 582 CGTTGGCAGAAATTTGCCAAACTACTTCCAGGAAGAGCTGATATCTATCAAAAATCAT 641

QY 111 TTPArgThrArgTleGlnLysHsileYsglnAlaGluAsnPhelngln-----Gln 128  
DB 642 TGGAAATCTACTATGCCAAGAAAGTGGAACAGAGAGCGCATTTTACAAGATGAATMAA 701

QY 129 SerSerAsnAsnSerGluileAsnAspHisGln-----AlaSerThrSerHisValSer 146  
DB 702 TCAGAAAGATCTTCACTTAACCTTCAACACAACTTGTCAGACTATGATATGCAAA 761

QY 147 ThrMetAlaGluProMetGluMetTyrSerPro-----ProCysTyrGlnGlyMet 163  
DB 762 ACCCAG-----AATCAGTTTTCATATACCTGTTTCAGATCCCTGGTATCAGTATG 812

QY 164 -----LeuGluProPheSerThr-----GlnPhePro 172  
DB 813 TCACCTGAAGCAATGTATAGAACATGTTACAGCCCTACTTCCCTTTATTCAGCAACC 872

QY 173 ThrIleAsnProasp 177  
DB 873 TTCATTGATGATGAT 887

RESULT 10  
US-09-402-929-1  
Sequence 1, Application US/09402929  
Patent No. 6410825

GENERAL INFORMATION:  
APPLICANT: Temple University - Of The Commonwealth System of Higher Education  
APPLICANT: Toscani, Antonio  
APPLICANT: Hatton, Kimi  
APPLICANT: Reddy, E. P.  
TITLE OF INVENTION: A-myb NULL MUTANT TRANSGENIC ANIMALS AND  
TITLE OF INVENTION: USES THEREOF  
NUMBER OF SEQUENCES: 7  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: SEIDEL, GONDA, LAVORGNA & MONACO, P.C.  
STREET: Suite 1800 Two Penn Center Plaza  
CITY: Philadelphia  
STATE: PA  
COUNTRY: U.S.A.  
ZIP: 19102

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/402,929  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/US98/06896  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Monaco, Daniel A.  
REGISTRATION NUMBER: 30,480  
REFERENCE/DOCKET NUMBER: 6056-214 PC  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (215) 568-8383  
TELEFAX: (215) 568-5549  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 3602 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-09-402-929-1

Alignment Scores:	
Pred. No.:	1.92e-25
Score:	291.50
Percent Similarity:	59.49%
Best local Similarity:	40.51%
Query Match:	25.62%
DB:	4
Length:	3602
Matches:	64
Conservative:	30
Mismatches:	51
Indels:	13
Gaps:	6

QY	12	ASPRGIVLVALRGLVSLGSLYPRCTPRHMETCIGLUSPRLLEULEWELSNTRY	31
Db	500	AATCCAGAAATTCGATAAAGGTCCTTTGGACGTAAAGAAAGATCAGAGGTTATTGAATTA	555
QY	32	ILEALASNHGILGVLGULGYVALITRPSNRSERLEUALALYSALALACLYLEULYS---	50
Db	560	GTTCAGAAATATGCGGCCCAAAAAGGGTCCTTATATTCACAAA-----CATTTAAAGCA	613
QY	51	ARGASNGILVYSSERCYSARGLLEUARGTRIPLEUASNTRYLEUARGPROASPVALARG	70
Db	614	AGAAATGGCAACAGCAGTGCACGAAAGATGGCACATCACTTCAACCTTAAGTTGAGAGAG	677
QY	71	GLYSANILETHTRPROGLUGLUGLNULEUULEWELSNTRYLEULALSLALSTRIPLYASN	90
Db	674	TCCTTCCTGCAGACAGAAAGAAAGACAGCATCATATATGAACACACACAGCGCTGGGAAC	733
QY	91	ARGTRPSERLYSLALALALYHISLEUPROGLYARGTHRASPASNGULILEYASNTYR	111
Db	734	CGTTGGGCCGAGAGCTGTAAGCTACTTCCTCGGAAGGCTGATATATCTTCAAAAATATAT	793
QY	111	TRPARGTHRARGILEGLINLHSHISLILEYSGINLAGLUSNPHEGLIN-----GLN	128
Db	794	TGGAAATTCATCCATCGAAGAAAGAAAGTGGAAACGAGGGCGATTTACAAACATGCAATAAA	853
QY	129	SERSEASNAHSERGLUILEASNAPSISGLN-----ALASERTHRSEHISVALSER	146
Db	854	TCGAGACGGGCTTCATCAAAAACATTCACACAAACCTTTGCGACTATGGACCATTTTGCAG	913
QY	147	THMETALAGLUPROMEGGLUWELTYSERP-----PROGYSITRYN	161
Db	914	ACCCAG-----AATCAGTTTACATTCCTCTTCAGATCCCTGGGATNTAG	958

```

? ATTORNEY/AGENT INFORMATION:
? NAME: Monaco, Daniel A.
? REGISTRATION NUMBER: 30,480
? REFERENCE/DOCKET NUMBER: 6056-214 PC
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: (215) 568-8383
? TELEFAX: (215) 568-5549
? INFORMATION FOR SEQ ID NO: 4:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 6775 base pairs
? TYPE: nucleic acid
? STRANDEDNESS: single
? TOPOLOGY: linear
US-09-402-929-4

```

```

Oy 12 AsprpogivalaIaglysglycylprotrpIthmetgyliaIasleuIleuehtksntr 31
Db 3749 AATCCAGATATGATAAAGGGCTCTTGGACTAAGAGAAAGATCAGAGGTTATGTAATTA 3808
Oy 32 ILeaIasnhtsgIyglugIyValITTPasnSerleuAlaIaIaIagIyIeulys--- 50
Db 3809 GTTTCAGAAATATGGGCCCAAAAAGGCTGCTTTTAATGGCAAAA-----CATTTAAAGCA 3862
Oy 51 ArgasnIyIysSerCysArgIeulAqtrpIeulAsnIyIeulArgProasrValArgArg 70
Db 3863 AGAATAGGCACAGACGATGCAAGAAAGATGGCGACAAATCAGCTTAACCCGTGAAGTGAAGAG 3922
Oy 71 GlyasnIlethrProglugIugIuIeulIeuerGlyIeulhIaIaIyIstPcIyasn 90
Db 3923 TCTTCTCGGACACAGAAAGAAAGACAGAGATCATATGAAAGCACACAAAGGCGCTGGGAAC 3982
Oy 91 ArgtrpserIyIleAlaIyIshIeulProgluYArgTrhAspAsclIyIyIysntr 110
Db 3983 CGTTGGGCGCGAGATGCTTAAGTACTTCTCGGAAGGAGCTGATAATCTTATCAAAATATAT 4042
Oy 111 TrpArgTrhArgIlegluIyIshIeulYsgIuIaIaIgluAsnRhegluIin-----glu 128
Db 4043 TGGAAATTCATCCACATGCCAAGAAAGAGTGGAAACAGAGGCGATTTTCAAGATGGATTAAGA 4102
Oy 129 SerSerAsnAsnSerGluIleAsnAspHisgluI-----AlaSerThiSerHisValSer 146
Db 4103 TCAAGAGCGGCTTCATCAAAAATCTCAACACAAACCTTGGCAGCATAGGACCATTTGCCA 4162
Oy 147 ThreAlaIagIuPrometGlymetYIySerPro-----ProCysIyIuIin 161
Db 4163 ACCCAG-----AATCAGTTTACATTCCTGTTTCAATCCCTGGGAGTATAG 4207

```





Query Match:	21.7%	Indels:	81
DB:	3	Gaps:	2
US-10-021-811-36 (1-206) x US-08-750-357-1 (1-4059)			
QY 15 VALRgrylsGlyProTrrpThrMetGluGluAspLeuIleLeuMetAsnTyrrIleAlaAsn 34			
DB 1111 GTTAGAGAGAGGGCGGTGGAGACCAAGAGAGAGATGCTTGCCCGCTACCTCAAGGCC 1170			
QY 35 HIsGlyGluGlyValTrrpAsnSerLeuAlaLys----- 45			
DB 1171 CATGGCCAAAGCAATATGAGAGAAAGT- GCCCCAGAAAGCCGGTAAACTAGTACTTTT 1225			
QY 45 ----- 45			
DB 1230 TTATTTCATTTTGAGATCATATATATACCCCGAGGCCAAGACCGAGAGCATCACTGT 1285			
QY 46 ---AlAlaGlyLeuLysAlaGAsnGlyLysSerCysArgLeuAlaArgTrpLeuAsnTyrrLeu 64			
DB 1290 GTGGGTACAGTTTGGCTCGGTGGGCAAGAGCTCGCGGCTGGGGCTGGAACACTCTC 1344			
QY 65 ArgProaspValArgArgGlyAsnIleThrProGluGluGluLeuIleMetGluLeu 84			
DB 1350 CGGCCCAACATCAGGCGCGCCCAACATCTCTACGAGAGAGAGATCTCATCATCCGCTC 1405			
QY 84 ----- 84			
DB 1410 CACAGGCTCTCGGCAACAGCTGTGTGAGAGTGCGCAGTGGCTGAGCTTATTACAGA 1465			
QY 84 ----- 84			
DB 1470 GCTGACGACGAGGAGCATGATCGATCGAGCGTCTGCGAATTCACTGTGCCGCTGCGCC 1529			
QY 85 -----HisAlaLysTrpGlyAsnArgTrpSerLysIle 95			
DB 1530 GTTGAGACGTGAGCTCATTTCAATATGTACATGCGCTGTGGCG- CGCAGGGGGTGGCTGATT 1588			
QY 96 AlAlaLysIleLeuProGlyArgTrpAspAsnGluIleLysAsnTyrrTrpArgThrArgIle 115			
DB 1589 GCAGGACGCGTGGCTGGCCGCAACAGCATGAAATCAAGAACTACTGCAACGACGCTG 1648			
RESULT 15			
US-08-485-139-5			
: Sequence 5, Application US/08485139			
: Patent No. 5880331			
GENERAL INFORMATION:			
APPLICANT: KREBBERS, Enno			
APPLICANT: WILLIAMS, Mark			
APPLICANT: LIEHMANS, Jan			
TITLE OF INVENTION: USE OF ANTHOCYANIN GENES TO MAINTAIN			
TITLE OF INVENTION: MALE STERILE PLANTS			
NUMBER OF SEQUENCES: 6			
CORRESPONDENCE ADDRESS:			
ADDRESSEE: Burns, Doane, Swecker & Mathis			
STREET: P.O. Box 1404			
CITY: Alexandria			
STATE: Virginia			
COUNTRY: United States			
ZIP: 22313-1404			
COMPUTER READABLE FORM:			
MEDIUM TYPE: Floppy disk			
COMPUTER: IBM PC compatible			
OPERATING SYSTEM: PC-DOS/MS-DOS			
SOFTWARE: Patent In Release #1.0, Version #1.25			
CURRENT APPLICATION DATA:			
APPLICATION NUMBER: US/08/485.139			
FILING DATE: 07-JUN-1995			
CLASSIFICATION: 800			
ATTORNEY/AGENT INFORMATION:			
NAME: McGowan, Malcolm K.			
REGISTRATION NUMBER: 39,300			
REFERENCE/DOCKET NUMBER: 010830-096			
TELECOMMUNICATION INFORMATION:			

TELEPHONE: (703) 836-6620  
TELEFAX: (703) 836-2021  
INFORMATION FOR SEQ ID NO: 5:

```

; OTHER INFORMATION: /note="TGCAG (in C1) which in C1-S allele is
; OTHER INFORMATION: replaced with TTAGG"
US-08-485-139-5

```

QY	15	ValArGysalIyProtrPrHrMetGluGluIuaSPleuIleuMeFaSnTyIlelaaSn	34
		:	
Db	1060	GTTbAGAGAGAGGGCGGTGGACGACGACAGGAGGACGATGCTTGCGCGCTACGTCAAGCC	1119
QY	35	HisSlYgluIyValTrpAsnSerIleuAlaIys-	45
Db	1120	CATGGCGAGAGGCAAAATGGAGGGAGAGT-GCCCCAGAAAGCGGTAAAAATAGACTAGTCTTT	1178
QY	45	-----	45
Db	1179	TTATTCATTTTGGATCATATATATACCCCGGAGGCAAGCCGAGAGACATCACTGT	1238
QY	46	---AlaAlaSlYleuIysArGbaEngIyIysSerCysArGleuArGTrPleuAsnTrIeu	64
Db	1239	GTGGGTGCAGGTTTGGGTGGTCCGGCAAGACTCCGGCGTGGGTGCACTACATCC	1298
QY	65	ArgProAspPalaIrrArgTrpGlyAsnIleThrProGluGluIleuIleuIleuMeGluIleu	84
		:	
Db	1299	CGGCCAACATCAGGCGCGGCAACATCTCTACGAGAGAGAGATCTATCATTCGCGCTC	1358
QY	84	-----	84
Db	1359	CACAGGCTCTCGGCAACAGTCTGTGCAGTGGCCAGTGGTGGCTAGCTATTATACAGA	1418
QY	84	-----	84
Db	1419	GCTGACGACGAGCCGATCGATCGAGCGCTGCTGCCGAATTCAATCGTTCCGGTGTGGCC	1478
QY	85	-----HisAlaIysTrpGlyAsnArgTrPserIysIle	95
Db	1479	GTGTGAGAGTGAAGTCATTCATATATATACATGCGTGTGGCG-CGAGAGTGGTGGCTGATT	1537
QY	96	AlaAlaIysHisIeuProGlyArGTrHrAspAsnGluIleIysAsnTyTrPArGTrHrArgIle	115
Db	1538	GCAGGCGAGGTGCTGGCCGCAACAGCATGAAATCAAAACTACTGTGAAACGACACGCTG	1597



GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame\_plus.p2n model

Run on: February 18, 2003, 09:25:50 ; Search time 220 Seconds  
(without alignments)  
2108.689 Million cell updates/sec

Title: US-10-021-811-36  
Perfect score: 1138  
Sequence: 1 MDRKQCKTSODPEYRKGPW.....NNINYSWSDSMQMLNGD 206

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 2185239 seqs, 112599159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL-frame-p2n.model -DEV-x1h  
-O-/cgn2.1/USPTO.spool/US10021811/runat.12022003.090133.9544/app.query.fasta.1.391  
-Db=N\_Geneseq.101002 -OPMT=firstap -SUFFIX=p2n.rng -MINMATCH=0.1 -DOOCT=0  
-LOOPEXT=0 -UNITS-bits -START=1 -END=1 -MATRIX-Biosum62 -TRANS-human40.cdi  
-LIST=45 -DOCALLIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=ptc -NORM-ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US10021811.ecgn.1.125.etrnat.12022003.090133.9544 -NCPU=6 -ICPU=3  
-NO\_XLPXY -NO\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -LONGLOG -DEV TIMEDOUT=120  
-MARN TIMEDOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N\_Geneseq\_101002.\*  
1: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:\*  
2: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:\*  
3: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:\*  
4: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:\*  
5: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:\*  
6: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:\*  
7: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:\*  
8: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:\*  
9: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:\*  
10: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:\*  
11: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:\*  
12: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:\*  
13: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:\*  
14: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:\*  
15: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:\*  
16: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:\*  
17: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:\*  
18: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:\*  
19: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:\*  
20: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:\*  
21: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:\*  
22: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:\*  
23: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:\*  
24: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	759.5	66.7	862	21	AAC57336	Eucalyptus grandis
2	671	59.0	910	21	AAC40587	Arabidopsis thalia
3	667	58.6	908	21	AAC48506	Arabidopsis thalia
4	625.5	55.0	907	21	AAC57337	Eucalyptus grandis
5	585	51.4	621	21	AAC41714	Arabidopsis thalia
6	585	51.4	1127	21	AAC37953	Arabidopsis thalia
7	540	47.5	673	21	AAC56197	Eucalyptus grandis
8	539	47.4	1137	22	AAD05784	Arabidopsis thalia
9	523.5	46.0	981	24	ABK65364	Arabidopsis thalia
10	510	44.8	972	24	ABK65364	Arabidopsis thalia
11	495	43.5	626	22	AAH87724	Pepermint plant o
12	478	42.0	524	21	AAC57194	Eucalyptus grandis
13	475	41.7	837	20	AAC25572	Arabidopsis thalia
14	459	40.3	959	21	AAC44894	Arabidopsis thalia
15	421	37.0	389	21	AAC56152	Eucalyptus grandis
16	421	37.0	417	21	AAC56774	Eucalyptus grandis
17	410.5	36.1	516	21	AAC56352	Pinus radiata tran
18	410.5	36.1	516	21	AAC56457	Pinus radiata tran
19	406.5	35.7	1384	21	AAC42228	Arabidopsis thalia
20	406	35.7	660	24	ABK65183	Arabidopsis thalia
21	405.5	35.6	542	21	AAC56434	Pinus radiata tran
22	404.5	35.5	389	21	AAC56837	Pinus radiata tran
23	404	35.5	1081	22	AAFP0596	Cotton transcript
24	402.5	35.4	1820	21	AAZ46048	CDNA sequence of a
25	402.5	35.4	1820	22	AAH28336	Nucleotide sequenc
26	400.5	35.2	774	21	AAC42662	Arabidopsis thalia
27	397.5	34.9	1209	22	AAD05748	Arabidopsis thalia
28	397.5	34.9	1209	24	ABK65181	Arabidopsis thalia
29	396.5	34.8	1006	22	AAFP0594	Cotton transcript
30	396	34.8	2220	18	AAC60744	Barley gibberellin
31	392	34.4	1218	21	AAC57346	Pinus radiata tran
32	391.5	34.4	612	24	ABK65317	Arabidopsis thalia
33	390.5	34.3	800	24	ABK65247	Arabidopsis thalia
34	390	34.3	526	21	AAC57245	Pinus radiata tran
35	387	34.0	1369	22	AAD05764	Arabidopsis thalia
36	386.5	34.0	1825	21	AAC44207	Arabidopsis thalia
37	386	33.9	2032	18	AAT60745	Arabidopsis thalia
38	385	33.8	1062	21	AAC43611	Rice gibberellin-r
39	385	33.8	1087	22	AAD06456	Arabidopsis thalia
40	384	33.7	1081	24	ABK65248	Arabidopsis thalia
41	383.5	33.7	949	22	AAFP5188	Nucleotide sequenc
42	383.5	33.7	2521	20	AAZ55571	Arabidopsis thalia
43	383	33.7	573	21	AAC57325	Pinus radiata tran
44	382.5	33.6	843	21	AAC43054	Arabidopsis thalia
45	382.5	33.6	1044	24	ABN98399	Arabidopsis thalia

## ALIGNMENTS

RESULT 1  
AAC57336  
ID AAC57336 standard; DNA; 862 BP.

AC AAC57336;

XX 25-JAN-2001 (first entry)

DE Eucalyptus grandis transcription factor DNA sequence #773.

XX Plant; transcription factor; gene expression; eucalyptus; pine; acacia;

KW poplar; sweetgum; teak; mahogany; bZIP; G-box binding factor;

KW basic helix-loop-helix zipper; homeotic; homeodomain; homeobox; MADS;

KW homeodomain zipper; LIM domain; AP2; ERBS; zinc finger domain;

XX type 2 Cys2His2; CCAAT box element; MYB; ss.

XX Eucalyptus grandis.

OS

PN WO2000053724-A2.  
XX  
PD 14-SEP-2000.  
XX  
PT 09-MAR-2000; 2000WO-US06112.  
PF  
PR 11-MAR-1999; 99US-0266513.  
PR 18-AUG-1999; 99US-0149485.  
PA (GENE-) GENESIS RES & DEV CORP LTD.  
XX (FLET-) FLETCHER CHALLENGE FORESTS LTD.  
XX  
PI Wood M., McGrath A., Shenk MA., Glenn M.  
DR WPI: 2000-579369/54.  
XX  
PT New isolated polynucleotide encoding a plant transcription factor for  
PT producing a plant e.g. a woody plant, preferably eucalyptus or pine,  
PT having modified gene expression or modified activity of a polypeptide  
-  
PS  
PS Claim 1; Page 649; 747pp; English.

The present invention relates to novel plant transcription factors from  
CC Eucalyptus grandis or pinus radiata. The present sequence is the coding  
CC sequence for one such transcription factor. The transcription factor may  
CC be used to produce a plant having modified gene expression such as a  
CC woody plant e.g. a eucalyptus, pine, acacia, poplar, sweetgum, teak, or  
CC many other species or to modify the activity of a polypeptide in a plant.  
CC The transcription factors of the present invention are members from the  
CC following families of regulatory proteins: bzip, bzip family of G-box  
CC binding factors, basic helix-loop-helix zipper,  
CC homeotic/homeodomain/homeobox/MADS, homeodomain zipper, LIM domain, AP2  
CC and ERBs, zinc finger domains of type 2 Cys2His2, CCAT box elements  
CC and XRBs.  
XX

Sequence 862 BP; 268 A; 187 C; 213 G; 194 T; 0 other:

Alignment Scores:  
Pred. No.: 1.7e-74 Length: 862  
Score: 759.50 Matches: 152  
Percent Similarity: 76.30% Conservative: 9  
Best local Similarity: 72.04% Mismatches: 29  
Query Match: 66.74% Indels: 21  
DB: 21 Gaps: 5

US-10-021-811-36 (1-206) x AAC57336 (1-862)

Gy 1 Metasplyls-----GlnGInCylslystHsrSerGlnAspProGluValArgLysGly 18  
Db 31 ATGGCAAGAAAGCCAGACGACGACAGTGTGAATGCCAAGAATGTCGAGGAGAAAAGGC 90  
Gy 19 ProTrpThrMetGluGluAspLeuIleUeuMeLsnAnryrIlleaAaSnHslsGlyGly 38  
Db 91 CGGTGGACGATGGAAGAGATCTCATCTCTCATCAACTATAGCGATCACGGGCAAGGC 150  
Gy 39 ValTrpAsnserLeuAlaLysAlaAlaGlyLeuUlysArgrnsnglyLysSeryAsrGlu 58  
Db 151 AGTTGGAACCTCCCTTAAGCCAAGCTGCTGCTTAATAAACGTACCAGGAAGATTGTGGCTC 210  
Gy 59 ArgTrpleuasnrYrLeuaRPrCoaspVlaIrgArGlLysanllerrProGluGln 78  
Db 211 CGGTGGCTGAACATCTCTGCGACCCGACGCTCCGGAGAGCAACATCACTACTGAGAGCAG 270  
Gy 79 LeuLeuIleMetGluLeuHIsAlaLysTrpGlyUsnaRGrPsrSerylIleaLalysHIs 98  
Db 271 CTCCTGATCATGGAACGTGCATGCCCAAGTGGGGAAACAAGTGTCTMAATTGCCAAGCAT 330  
Gy 99 LeuProGlyArgTrpAspAsnGluIleLysAsnrYrTrpArGrHrArgIleGlnLysHIs 118  
Db 331 CTTCCTGGAAAGACTGACATGAGATTAAGAAGACTTCTGGAGAGCTGAATCCAAAAAGCAC 390  
Gy 119 IleLysGlnAlaGluAsnPheGlnGlnInSerSerAsnAsnserGluIleAsnAspHIs 138

[illegible]

PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 14-JUN-1999; 99US-0138847.  
PR 16-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140699.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142053.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142927.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144003.  
PR 16-JUL-1999; 99US-0144086.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144684.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.

PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151066.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158232.  
PR 12-OCT-1999; 99US-0158369.  
PR 13-OCT-1999; 99US-0159294.  
PR 13-OCT-1999; 99US-0159295.  
PR 14-OCT-1999; 99US-0159328.  
PR 14-OCT-1999; 99US-0159330.  
PR 14-OCT-1999; 99US-0159331.  
PR 14-OCT-1999; 99US-0159637.  
PR 14-OCT-1999; 99US-0159638.  
PR 18-OCT-1999; 99US-0159584.  
PR 21-OCT-1999; 99US-0160741.  
PR 21-OCT-1999; 99US-0160767.  
PR 21-OCT-1999; 99US-0160768.  
PR 21-OCT-1999; 99US-0160770.  
PR 21-OCT-1999; 99US-0160814.  
PR 21-OCT-1999; 99US-0160815.  
PR 22-OCT-1999; 99US-0160980.  
PR 22-OCT-1999; 99US-0160981.  
PR 22-OCT-1999; 99US-0160989.  
PR 25-OCT-1999; 99US-0161404.  
PR 25-OCT-1999; 99US-0161405.  
PR 25-OCT-1999; 99US-0161406.  
PR 26-OCT-1999; 99US-0161359.  
PR 26-OCT-1999; 99US-0161360.  
PR 26-OCT-1999; 99US-0161361.  
PR 28-OCT-1999; 99US-0161920.  
PR 28-OCT-1999; 99US-0161922.



```

PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143622.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144337.
PR 19-JUL-1999; 99US-0144338.
PR 19-JUL-1999; 99US-0144339.
PR 19-JUL-1999; 99US-0144339.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 23-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 26-JUL-1999; 99US-0145226.
PR 27-JUL-1999; 99US-0145213.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 03-AUG-1999; 99US-0146389.
PR 04-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 06-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 26-AUG-1999; 99US-0150884.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151348.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.

```

```

PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161933.
PR 29-OCT-1999; 99US-0162142.

```

## Alignment Scores:

```

Pred. No.: 3,226-64
Score: 667.00
Percent Similarity: 72.43%
Best Local Similarity: 61.68%
Query Match: 58.61%
DB: 21
Gaps: 4

```

US-10-021-811-36 (1-206) x AAC48506 (1-908)

```

QY 1 MetAspLysLysGlnGlnCys-----LysThrSerGlnAspProGluValArgLysGly 18
DB 108 ATGAGAGAAAAGAGAAAGTGTGGTGTGATCGAGAGATGAGAGATGAGAGAAAGG 167
QY 19 ProTriPheMetGluGluAspLeuIleLeuMetAsnTyrlleAlaAsnHISGlyGluGly 38
DB 168 CCATGACGACATGAGAGAAAGATTTGATTCATCATTAATGCGCAATCATGCGAAAGCT 227
QY 39 ValTrpAsnSerLeuAlaIysAlaIagIleuLysAlaGlnGlyLysSerCysArgLeu 58
DB 228 GTTGGAACTCTCCGCAAAATCTGCGAGACTAAAGCAGCCGGAAGAAAGTTCGCGGCTC 287
QY 59 ArgTrpLeuAsnTyrlleuArgProAspValArgArgGlyAsnIleThrProGluGluGln 78
DB 288 CGGTGGCTGACTACTCTCGACCTGATGTGCGAGCGGGAATATATACACCGAGAGAAAG 347
QY 79 LeuLeuIleMetGluLeuMetAlaIysAlaIysArgGlyAsnArgTrpSerLysIleAlaLysHis 98
DB 348 CTCACCATCAWGCAGACTCATGCAAAATGCGGAATAGTGCTCAAAAATTCGCAAAAGCAT 407
QY 99 LeuProGlyArgThrAspAsnGluIleLysAsnTyrlleArgIleGlnLysHis 118

```



KW Hybridisation assay; genetic mapping; gene expression control;  
KW protein identification; signal transduction pathway;  
XX metabolic pathway; promoter; termination sequence; ss.  
OS Arabidopsis thaliana.  
XX  
PN EP1033405-A2.  
XX  
PD 06-SEP-2000.  
XX  
XX 25-FEB-2000; 2000EP-0301439.  
XX  
PR 25-FEB-1999; 99US-0121825.  
PR 05-MAR-1999; 99US-0122180.  
PR 09-MAR-1999; 99US-0123348.  
PR 23-MAR-1999; 99US-0125788.  
PR 25-MAR-1999; 99US-0126264.  
PR 29-MAR-1999; 99US-0126785.  
PR 01-APR-1999; 99US-0127462.  
PR 06-APR-1999; 99US-0128234.  
PR 08-APR-1999; 99US-0128714.  
PR 16-APR-1999; 99US-0128845.  
PR 19-APR-1999; 99US-0130077.  
PR 21-APR-1999; 99US-0130449.  
PR 23-APR-1999; 99US-0130510.  
PR 23-APR-1999; 99US-0130891.  
PR 28-APR-1999; 99US-0131449.  
PR 30-APR-1999; 99US-0132048.  
PR 30-APR-1999; 99US-0132407.  
PR 04-MAY-1999; 99US-0132484.  
PR 05-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.  
PR 06-MAY-1999; 99US-0132487.  
PR 07-MAY-1999; 99US-0132863.  
PR 11-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139434.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-013950.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140053.  
  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142927.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144326.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 08-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 23-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.

```
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159320.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 26-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161993.
PR 28-OCT-1999; 99US-0162142.
PR 29-OCT-1999; 99US-0162142.

Alignment Scores:
Pred. No.: 2,29e-55
Score: 585.00
Percent Similarity: 65.14%
Best local Similarity: 55.50%
Query Match: 51.41%
DB: 21
Gaps: 6
```

US-10-021-811-36 (1-206) x AAC1714 (1-621)

```
QY 1 MetaspLysGlnGlnCysLys-----ThrSerGlnAspProGlu----- 14
DB 13 ATGAAGAAGAAAGGAGGAGCAATACGTCACAGAAAGAAAGAAAGAAC 72
QY 15 ValAlaGlySerGlyProThrThrMetGlnGlnAspLeuIleLeuMetAsnTyIleAlaAsn 34
DB 73 GTGAAGAAAGACCTTGACATATGAAGAAATTTTCATCCCTTTAATTATCATCTTAAT 132
QY 35 HisGlyGlnGlyValTyrPAsnSerLeuAlaLysAlaGlyLeuLysAlaGlyLys 54
DB 133 CATGGTGAAGGCTTTGGAACTGTGCGCAAGGCTCTGTGCTAAGAAACGTACTGGAAA 192
QY 55 SerCysArgLeuArgTyrPLeuAsnTyrLeuArgProAspValArgArgGlyAsnIleThr 74
DB 193 AGTTGTGGCTCGGCTGGCTGAACATATCTCGACCAAGATGTGGCGGAGGAGACATAAC 252
```

```
QY 75 ProGlnGlnGlnLeuLeuIleMetGlnLeuHisAlaLysTyrPAsnArgTyrSerLys 94
DB 253 GCAGAAAGACAGACCTTTATCATTCATGCTTAAGCTTGGAACAGGTGTCAGAG 312
QY 95 IleAlaLysHisLeuProGlyArgThrAspAsnGluIleLysAsnTyrTyrPArgThrArg 114
DB 313 ATTGGGAAGCATCTTCGCGGAAGAACGACAAAGATAAAGAACTCTGTGAGGACAAAG 372
QY 115 IleGlnLysHisIleLysGlnAlaGlu-----AspPheGlnGlnGlnSer 130
DB 373 ATTTCAGACACACATCATAGTGTATCGAATAATGATGAAATCATCAATCATTTCTTCG 432
QY 131 AsnAsnSerGluIleAsn-----AspHisGlnAlaSerThrSerHisValSerThrMet 148
DB 433 GGAAACTCAGACAGCTCGCGGATGACGACGACGACGACCTCGGCAAGCCATGACACAG 492
QY 149 AlaGluProMetGlnMetLysSerProProCysTyrGlnGlyMetLeuGluProPheSer 168
DB 493 GCTGAGAGCTTCTCT-----CAGGCGAAGAG 519
QY 169 ThrGlnPheProThrIleAsnProAspGlnSerSerCysTyrAsnAspAsnAsn 188
DB 520 ACGACGTTTAATGTGTGTC-----GAAACACAGTCAAC 552
QY 189 IleAsnTyrTrpSerMetGlnAspSerTrpSerMetGlnLeuLeuAsnGlyAsp 206
DB 553 GAGAAATTAAGTGAAGCTGAGATCTGTGCGCCGTCACCTTGAATGTGAC 606

RESULT 6
AAC37953
ID AAC37953 standard; DNA; 1127 BP.
XX
AC AAC37953;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 19258.
XX
KW Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway;
KW metabolic pathway; promoter; termination sequence; ss.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0128845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
```



PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137502.  
PR 04-JUN-1999; 99US-0137724.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139493.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 21-JUN-1999; 99US-0139767.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139869.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140699.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142927.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 20-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 03-AUG-1999; 99US-0146389.  
PR 04-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 05-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 06-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151308.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158232.  
PR 12-OCT-1999; 99US-0158369.  
PR 13-OCT-1999; 99US-0159293.  
PR 13-OCT-1999; 99US-0159294.  
PR 13-OCT-1999; 99US-0159295.  
PR 14-OCT-1999; 99US-0159329.  
PR 14-OCT-1999; 99US-0159330.  
PR 14-OCT-1999; 99US-0159331.  
PR 14-OCT-1999; 99US-0159637.  
PR 14-OCT-1999; 99US-0159638.  
PR 18-OCT-1999; 99US-0159584.  
PR 21-OCT-1999; 99US-0160741.  
PR 21-OCT-1999; 99US-0160767.  
PR 21-OCT-1999; 99US-0160768.  
PR 21-OCT-1999; 99US-0160770.  
PR 21-OCT-1999; 99US-0160814.  
PR 21-OCT-1999; 99US-0160815.  
PR 22-OCT-1999; 99US-0160980.  
PR 22-OCT-1999; 99US-0160981.

PR 22-OCT-1999; 99US-0160989.  
 PR 25-OCT-1999; 99US-0161404.  
 PR 25-OCT-1999; 99US-0161405.  
 PR 25-OCT-1999; 99US-0161406.  
 PR 26-OCT-1999; 99US-0161359.  
 PR 26-OCT-1999; 99US-0161360.  
 PR 26-OCT-1999; 99US-0161361.  
 PR 28-OCT-1999; 99US-0161920.  
 PR 28-OCT-1999; 99US-0161921.  
 PR 28-OCT-1999; 99US-0161923.  
 PR 29-OCT-1999; 99US-0162142.

## Alignment Scores:

Pred. No.:	5,26e-55	Length:	1127
Score:	585.00	Matches:	121
Percent Similarity:	65.148	Conservative:	21
Best Local Similarity:	55.508	Mismatches:	44
Query Match:	51.418	Indels:	32
DB:	21	Gaps:	6

US-10-021-811-36 (1-206) x AAC37953 (1-1127)

QY 1 MetAspLysLysGlnGlnCysLys-----ThrSerGlnAspProGlu----- 14  
 Db 286 ATGAAGAAGAAAGGAGAGTGAAGACACAAATTAACGTACAGAAAGAAAGAAAGAACAA 345  
 QY 15 ValArgLysGlyProTrpThrMetGluGluAspLeuIleLeuMetAsnTrpIleAlaAsn 34  
 Db 346 GTGGAAGAAAGACCTTGACATCGAAGAGATTTCCTCTTTAATTACATTCCTTAAT 405  
 QY 35 HisGlyGlnGlyValTrpAsnSerLeuAlaLysAlaAlaGlyLeuLysAlaGAsnGlyLys 54  
 Db 406 CATGCTGAAGTCTTTTGAACTCTGTGCCAAGCCCTGTGTAAACCTACTGCGAAA 465  
 QY 55 SerCysArgLeuArgTrpLeuAsnTrpLeuArgProAspValArgArgGlyAsnIleThr 74  
 Db 466 AGTTGTGGCTCCGCGTGGCTGAACTATCTCGACCAATGTGCGGCAAGCAATTAAC 525  
 QY 75 ProGlnGlnGlnLeuLeuIleMetGluLeuHisAlaLysTrpGlyAsnArgTrpSerLys 94  
 Db 526 ACAGAAAGAACGCTTTTGATCATCTCACGCTTCATGCTGAAGCTTGGAACAGGTGCGAAG 585  
 QY 95 IleAlaLysHisLeuProGlyArgTrpAspAsnGluIleLysAsnTrpTrpArgThrArg 114  
 Db 586 ATTCGGAAGCATCTTCGCGAAGAACGACACAGATTAAGAACTTCTGAGAGACAAAG 645  
 QY 115 IleGlnLysHisIleLysGlnAlaGlu-----AsnProGlnGlnGlnSerSer 130  
 Db 646 ATTCAGAGACACATGAAGTGTATCGAAGAAATATGATGAATCATCAATCATATTGTTGC 705  
 QY 131 AsnAsnSerGluIleAsn-----AspHisGlnAlaSerThrSerHisValSerThrMet 148  
 Db 706 CGAAGCTCAGACGCTCGGGATGACGACGACGACGACGCTCCGCAAGCCATAGACACG 765  
 QY 149 AlaGluProMetGluMetLysSerProCysTrpGlnGlyMetLeuGluProPheSer 168  
 Db 766 GCTGAGACCTTCTCT-----CAAGCGCAAGACG 792  
 QY 169 ThrGlnPheProThrIleAsnProAspGlnSerSerCysTrpAsnAspAsnAsn 188  
 Db 793 ACGACGTTAATGTGGT-----GAACAACAGCTCAAC 825  
 QY 189 IleAsnTrpTrpSerMetGluAspSerTrpSerMetGlnLeuLeuAsnGlyAsp 206  
 Db 826 GAGAAATTAATCGAAGTTGAGATCTGTGCGCTCCACTGTTAATGTGTAC 879

## RESULT 7

AAC56197  
 ID AAC56197 standard; DNA; 673 BP.  
 AC AAC56197;  
 XX  
 XX 25-JAN-2001 (first entry)

XX Eucalyptus grandis transcription factor DNA sequence #328.  
 DE  
 XX

plant; transcription factor; gene expression; eucalyptus; pine; acacia;  
 poplar; sweetgum; teak; mahogany; bzlp; G-box binding factor;  
 basic helix-loop-helix zipper; homeotic; homeodomain; homeobox; MADS;  
 homeodomain zipper; LIM domain; AP2; EREBS; zinc finger domain;  
 type 2 Cys2His2; CCAAT box element; MYB; ss.

Eucalyptus grandis.

WO200053724-A2.

14-SEP-2000.

09-MAR-2000; 2000WO-US06112.

11-MAR-1999; 99US-0266513.

18-AUG-1999; 99US-0149485.

(GENE-) GENESIS RES & DEV CORP LTD.  
 (FLET-) FLETCHER CHALLENGE FORESTS LTD.

Wood M, McGrath A, Sheak MA, Glenn M;

WPI: 2000-579369/54.

New isolated polynucleotide encoding a plant transcription factor for  
 producing a plant e.g. a woody plant, preferably eucalyptus or pine,  
 having modified gene expression or modified activity of a polypeptide

Claim 1; Page 131; 747pp; English.

The present invention relates to novel plant transcription factors from  
 Eucalyptus grandis or Pinus radiata. The present sequence is the coding  
 sequence for one such transcription factor. The transcription factor may  
 be used to produce a plant having modified gene expression such as a  
 woody plant e.g. a eucalyptus, pine, acacia, poplar, sweetgum, teak, or  
 mahogany species or to modify the activity of a polypeptide in a plant.  
 The transcription factors of the present invention are members from the  
 following families of regulatory proteins: bzlp, bzlp family of G-box  
 binding factors, basic helix-loop-helix zipper,  
 homeotic/homeodomain/homeobox/MADS, homeodomain zipper, LIM domain, AP2  
 and EREBS, zinc finger domains of type 2 Cys2His2, CCAAT box elements  
 and MYB.

Sequence 673 BP; 143 A; 227 C; 196 G; 107 T; 0 other;

## Alignment Scores:

Pred. No.:	2.47e-50	Length:	673
Score:	540.00	Matches:	113
Percent Similarity:	61.998	Conservative:	24
Best Local Similarity:	51.138	Mismatches:	36
Query Match:	47.458	Indels:	48
DB:	21	Gaps:	6

US-10-021-811-36 (1-206) x AAC56197 (1-673)

QY 17 LysGlyProTrpThrMetGluGluAspLeuIleLeuMetAsnTrpIleAlaAsnHisGly 36  
 Db 2 AGAGTCCCTGGACGCTGAGAGAGACGCTACCTGTAATTAATTAATTAATTAATTAATTA 61  
 QY 37 GluGlyValTrpAsnSerLeuAlaLysAlaAlaGlyLeuLysArgAsnGlyLysSerCys 56  
 Db 62 GAAGGACGCTGGAACTCCCTCGCCGACGACGAGGTTGAAACGACGCAAGGAGGTTGC 121  
 QY 57 ArgLeuArgTrpLeuAsnTrpLeuArgProAspValArgArgGlyAsnIleThrProGlu 76  
 Db 122 CGGCTCGGCGGTGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 181  
 QY 77 GluGlnLeuLeuIleMetGluLeuHisAlaLysTrpGlyAsnArgTrpSerLysIleAla 96

```

Db 182 GAGCAGCTCTTGATCTCGAGCTCCATTCGCCGTGGGCAATTCAGTGTCCAGATCGCG 241
Qy 97 LysHisLeuProGlyArgThrAspAsnGluIleuYsaNtyrTrpArgIleGln 116
Db 242 CAACACTTGGCGGCGAGGACGACAAAGATCAAGAACTACTGGCCAGACCGGGTGCAG 301
Qy 117 LysHisIleuYsaGlnAlaGluAsnPhenGlnGlnSerSerAsnAsnSerGluIleAsn 136
Db 302 AAGCAGCGGAAGCAGCTC-----AAATGTGACGTCAACAGCAACAGTTCAAG 349
Qy 137 ASPHis-----GlnAlaSerThrSer 143
Db 350 GACGCGCATGAATACCTGTGATGCCAGGCTGGTCGAGAGATCCAAAGCCGCTCCGCC 409
Qy 144 HisValSerThr-----MetaIaGluPrometGluMetTyrSer 156
Db 410 TCTGTCTGACCGCTACTGTGCGCGCCGCCCATGGACGCCACCAACCAATGGCCACC 469
Qy 157 -----ProProCysTyrGlnGlyMetLeuGlu 165
Db 470 ACCGCGCATCCAAACATCGCGCGCATGGCTTTCGCCGCCGCGCATGGCGCGCGC 529
Qy 166 ProPheSer-----ThrGlnPheProThrIleAsnProAspGlnSerSerCys 181
Db 530 GACTTCAGGGCGGGCGAGTGAATGTGGCGCCAGCTACACACCAGCGAGAACTCCTGC 589
Qy 182 CysThrAsnAspAsnAsnAsnIleAsnTyrTrpSerMetGluAspSerTyrSerMetCln 201
Db 590 ACGACG-----GCCGCTTCCGACTCATCTGCGTGCAGC 622
Qy 202 Leu 202
Db 623 CTC 625

RESULT 8
AAD05784
ID AAD05784 standard; cDNA; 1137 BP.
AC AAD05784;
XX
DT 31-JUL-2001 (first entry)
XX
DE Arabidopsis thaliana transcription factor, GI324 cDNA.
XX
KW Transcription factor; biochemical characteristic; controlling element;
KW structural characteristic; developmental characteristic; gene therapy;
KW agricultural biotechnology; plant trait modification; ss.
XX
OS Arabidopsis thaliana.
XX
FH Key Location/Qualifiers
FT CDS 54..914
FT /*tag= a
FT /product= "Transcription factor, GI324"
XX
XX MO200136597-A1.
XX
PD 25-MAY-2001.
XX
PF 14-NOV-2000; 2000MO-US31344.
XX
XX 17-NOV-1999; 99US-0166228.
PR 17-APR-2000; 2000US-0197899.
PR 22-AUG-2000; 2000US-0227439.
XX
XX (MEND-) MENDEL BIOTECHNOLOGY INC.
PA (CRE/) CREELMAN R.
PA (YUG/) YU G.
PA (ADAM/) ADAM L.
PA (RIEC/) RIECHMANN J L.
PA (HEAR/) HEARD J.
PA (SAMA/) SAMAH R.
PA (PILG/) PILGRIM M.

```

```

PA (PINE/) PINEDA O.
PA (JIAN/) JIANG C.
XX
PI Creelman R, Yu G, Adam L, Riechmann JL, Heard J, Samaha R;
PI Pilgrim M, Pineda O, Jiang C;
XX
DR WPI; 2001-335999/35.
XX P-PSDB; AA01900.
PT Nucleic acids encoding plant transcription factor polypeptides, useful
PT for altering the biochemical characteristics of plants e.g. corn,
XX potato and cotton plants -
XX
PS Claim 4; Page 87-88; 127pp; English.
XX
CC The present sequence is Arabidopsis thaliana transcription factor,
CC GI324 cDNA. The transcription factor is used for altering a plant's
CC biochemical characteristics. The transcription factor may be used to
CC alter the structure and developmental characteristics of plants such as
CC soybean, wheat, corn, potato, cotton, rice, oilseed rape, sunflower,
CC alfalfa, sugar cane, turf, banana, blackberry, blueberry, strawberry,
CC raspberry, cantaloupe, carrot, cauliflower, coffee, cucumber, eggplant,
CC grapes, honey dew, lettuce, mango, melon, onion, papaya, peas, peppers,
CC pineapple, spinach, squash, sweet corn, tobacco, tomato, watermelon,
CC roseaceous fruits and/or vegetable brassicas. Transcription factors are
CC key controlling elements of biological pathways and altering expression
CC levels of 1 or more transcription factors can change entire biological
CC pathways in an organism. Therefore manipulating transcription factor
CC levels in plants offers great potential in agricultural biotechnology in
CC for modifying a plant's traits. Transcription factor cDNA is useful in
CC gene therapy.
XX
SQ Sequence 1137 BP; 378 A; 213 C; 233 G; 313 T; 0 other;

Alignment Scores:
Pred. No.: 6.63e-50 Length: 1137
Score: 539.00 Matches: 107
Percent Similarity: 66.18% Conservative: 28
Best Local Similarity: 52.45% Mismatches: 43
Query Match: 47.36% Indels: 26
DB: Gaps: 5

US-10-021-811-36 (1-206) x AAD05784 (1-1137)
Qy 1 MetaSplysLysGlnGlnCysIlyrThrSerGlnAspProGluValArgLysGlyProTrp 20
Db 66 ATGACAGAGAGAGAGAGAGCTTCAAGAAAGTGAAGTGAAGAACTAAGAGAGCGCTTGG 125
Qy 21 ThMetGluGluAspLeuIleuMetAsnTyrIleAlaAsnHisGlyGluIlyValTrp 40
Db 126 ACTTTCGAGGAACACACACTTTCACCAAAATTACATCTCCATACGCGTGGTGG 185
Qy 41 AsnSerLeuAlaIlyAlaIlyLeuIlyAsnGlyLysSerCysArgLeuArgTrp 60
Db 186 AATCAGGTGCAAAATGTCTGGCTAAAGAAAGTGGGAAAGTGTGATGATGGATGG 245
Qy 61 LeuAsnTyrLeuArgProAspValArgArgLysAsnIleThrProGluGlnIleuLeu 80
Db 246 TTGAAATTACTTGAACCGACATTAAGACGAGGAATCTTACCTCCACAAACAGCTTTTG 305
Qy 81 IleMetGluLeuHisAlaIlyStrpGlyAsnArgTrpSerLysIleAlaLysHisLeuPro 100
Db 306 ATCTTTCGAGCTTCACTTAATGCGGTAAATGAGTGTCCAAAGATTGGACAGTACTTGGCA 365
Qy 101 GlyArgThrAspAsnGluIleuYsaNtyrTrpArgIleGlnLysHisIleYsa 120
Db 366 GGAAGAACGATTAACGAGTCAAGAACTATTGGAGAACAGAGTTCAAAACAGCTGCT 425
Qy 121 GlnAlaGluAsnPhenGlnGlnSerSerAsnAsnSerGluIleAsnAspHisGlnAla 140
Db 426 CAACCTC---AACATCGAA-----TCTACACGCGCAACAGTTCTTGGAC---GCTGTT 470
Qy 141 SerThrHisValSerThrMetAlaGluPrometGluMetTyrSerProProCysTyr 160

```

DB 471 CGTGTGTTTGGTCCCTACATTCATGAGAGATGCAACAAACTCA----- 518  
QY 161 GlnclmethylleugluPropheserThglnPheProthrIleasnProaspGlnSer 180  
DB 519 -----TCCACTACTACTAT 536  
QY 181 CysCysThrAsnAspAsnAsnIleAsnTyrTrp-----SerMetGluAspSer 197  
DB 537 TGGTGTCCCAAAACAAACAAACAACTCTCTCTCTCTCTCTCTCTCAATCTCACGACTCT 596  
QY 198 TrpSerMetGln 201  
DB 597 TTAAGTATGCAA 608  
RESULT 9  
ABK65252  
ID ABK65252 standard; cDNA; 981 BP.  
XX  
XX ABK65252;  
AC  
XX 02-JUL-2002 (first entry)  
DT  
XX Arabidopsis cDNA encoding a transcription factor #104.  
DE  
XX Plant; ss; gene; transcription factor; transgenic;  
KW agriculture; metabolic chemical; environmental stress; drought;  
KW microbial disease resistance; herbicide resistance; seed yield;  
KW fruit; yield; growth rate; leaf senescence; flower senescence.  
XX Arabidopsis thaliana.  
OS  
XX W0201215675-A1.  
PN  
XX 28-FEB-2002.  
PD  
XX 22-AUG-2001; 2001WO-US26189.  
PE  
XX 22-AUG-2000; 2000US-227439P.  
PR 16-NOV-2000; 2000US-0713994.  
PR 16-APR-2001; 2001US-0837944.  
XX  
XX (MEND-) MENDEL BIOTECHNOLOGY INC.  
PA (PIV3/) PILGRIM M.  
PA (CRE3/) CREELMAN R.  
PA (DUB3/) DUBBEL A J.  
PA (HEAR/) HEARD J.  
PA (JIAN/) JIANG C.  
PA (KED2/) KEDDIE J.  
PA (ADAM/) ADAM L.  
PA (RATC/) RATCLIFF O.  
PA (REUB/) REUBER J L.  
PA (RIE2/) RIECHMANN J L.  
PA (YUG3/) YU G.  
PA (PINE/) PINEDA O.  
PI Pilgrim M, Creelman R, Dubell AJ, Heard J, Jiang C, Keddie J;  
PI Adam L, Ratcliff O, Reuber JL, Riechmann JL, Yu G, Pineda O;  
XX WPI: 2002-292022/33.  
XX P-PSDB: AA093066.  
XX  
XX An isolated or recombinant polynucleotide used to produce a transgenic  
PT plant -  
PS  
XX Claim 4; Page 443-445; 941pp; English.  
XX  
XX The invention relates to 1 of 232 isolated or recombinant polynucleotides  
CC encoding an Arabidopsis thaliana transcription factor, their variants,  
CC complements, fragments, or related polynucleotide with 31% to 95%  
CC sequence identity, where the plant possesses an altered trait as compared  
CC to a wild-type or reference plant, or the plant exhibits an altered  
CC phenotype as compared to a wild-type or reference plant, or the plant

CC exhibits ectopic expression or altered expression of one or more genes  
CC associated with a plant trait as compared to a wild plant. Also included  
CC are a transgenic plant comprising the polynucleotides, a computer  
CC readable medium having stored sequence information, and identifying a  
CC homologue sequence from a database comprising a plurality of known plant  
CC sequences comprising inputting sequence information selected from one of  
CC 464 fully defined sequences given in the specification. The isolated or  
CC recombinant polynucleotide is used for producing a plant having a  
CC modified trait, the method comprising selecting a polynucleotide that  
CC encodes a polypeptide or an antisense nucleic acid, inserting the  
CC polynucleotide or antisense nucleic acid into an expression vector,  
CC introducing the vector into a plant or a cell of a plant to overexpress  
CC the polypeptide or antisense nucleic acid, thereby producing a modified  
CC plant, and selecting for a modified trait (e.g. increased  
CC production of agriculturally useful proteins or metabolic chemicals,  
CC pest tolerance, environmental stress response (e.g. drought), microbial  
CC disease resistance, herbicide resistance, seed and fruit yield, growth  
CC rate, leaf and flower senescence and many other traits listed in the  
CC specification). The present sequence is one of the 232 polynucleotides  
CC encoding an A. thaliana transcription factor.  
XX  
SQ Sequence 981 BP; 319 A; 193 C; 235 G; 234 T; 0 other;  
XX  
Alignment Scores:  
Pred. No.: 2,81e-48 Length: 981  
Score: 523.50 Matches: 119  
Percent Similarity: 55.77% Conservative: 26  
Best local Similarity: 45.77% Mismatches: 32  
Query Match: 46.00% Indels: 83  
DB: 24 Gaps: 9  
US-10-021-811-36 (1-206) x ABK65252 (1-981)  
QY 8 LysThrSerGlnAspProGluValArgLysGlyProThrPheMetGluLysPheLeu 27  
DB 90 AAAGTGGAGAAAGAAATGAGCGCTAAGAGACAGCGCCATGGACGTTGAGAAATTTAGG 149  
QY 28 LeuMetAsnTyrIleAlaAsnHisGlyGluGlyValTTPAsnSerLeuAlaLysAla 47  
DB 150 CTCATCAATTACATTCTGTGTGATGAGGATGAGATGAGATCTCTCTCTGTTGGCGC 209  
QY 48 G1YleuLysArgAsnG1YlyssSerCysArgLeuArgTyrPheAsnTyrLeuArgProasp 67  
DB 210 GAAGTCAAAAGAGACCGGAAAGAAAGCTGCACAGCTTCGGTGGCTGACATGTCCGACACAGT 269  
QY 68 ValArgArgGlyAsnIleThrProGluGluGlnLeuLeuIleMetGluLeuHisAlaLys 87  
DB 270 GTGGCGCGTGGAAACATTAACCTCGAAGAAACAACCTGTGATCTGTGAACCTTCACACACGT 329  
QY 88 TrpGlyAsnArgTyrPheSerLysIleAlaLysHisLeuProGlyArgThrAspAsnGluLe 107  
DB 330 TGGGGCAATGATGGTCTAAGATTGCACAATATTTCAGGAAGAAAGGATTAACGAGATC 389  
QY 108 LysAsnTyrTrpArgThrArgIleGlnLysHisIleLysGlnAlaGlu-----Asn 124  
DB 390 AAAAATCATTTGGAGAACACAGCTGTCAAAAGCAATGCAAAACAGCTTAAGCGACGTGAAC 449  
QY 125 PheGlnGln----- 127  
DB 450 AGTCAACAATTTAAAGACACCATGAAGTATCTTTGGATGCTCGGCTGTAGAAAGATC 509  
QY 128 -----GlnSerSerAsn 132  
DB 510 CAAAGCCGCGTCATCGGGGTGTTCATGTCATCTTGGCGTCAACCACTCCATGATCG 569  
QY 133 SerGluIleAsnAspHisGlnAlaSerThrSerHisValSerThrMetAlaGluProMet 152  
DB 570 TTGCTGATCAACAACAAC-----AACACCAACAAGCTGATATTTGGCT----- 614  
QY 153 GluMetLysSerProProCysTyrGlnGlyMetLeuGluProPheSerThrGlnPhePro 172  
DB 615 TTAATGAGTAACCATGATGCTTAC----- 638

Oy	173	Thrileasnprrpspcinsersercyscthrasn-----	184
Db	639	---ATCAGCGGGGATTAATTCCACGCCGTGGCAGATCTCTCATTCAGATTGCAGAGTGT	695
Oy	185	-----AspaAnsqAnsnIle-----	189
Db	696	CAAOTGACTAGTAGAAGTGCGGAAGATTGGTCAGATGCAAAATTCACCAAAAAATG	755
Oy	190	-----AsatYtRrpserMetGluaSpserTrpserMetGlueLeuAnaclYasp	206
Db	756	ACATCGCGCGGAAATTAT-----ATGGATTAAT-----ACGACGTGCACTATTAAACGAGAT	803
RESULT 10			
ABK65364			
ID	ABK65364	standard; cDNA; 972 BP.	
XX	ABK65364:		
XX	02-JUL-2002	(first entry)	
De	Arabidopsis cDNA encoding a transcription factor #216.		
KW	Plant; ss: gene; transcription factor; transgenic;		
KM	agricultural; metabolic chemical; environmental stress; drought;		
KW	microbial disease resistance; herbicide resistance; seed yield;		
KM	fruit yield; growth rate; leaf senescence; flower senescence.		
XX			
OS	Arabidopsis thaliana.		
FN	WO200215675-A1.		
XX	28-FEB-2002.		
XX	22-AUG-2001; 2001WO-US26189.		
PR	22-AUG-2000; 2000US-227439P.		
PR	16-NOV-2000; 2000US-0713894.		
PR	16-APR-2001; 2001US-0837944.		
XX	(MENDEL) MENDEL BIOTECHNOLOGY INC.		
PA	(PIG/) PILGRIM M.		
PA	(CREE/) CREELMAN R.		
PA	(DUBE/) DUBELL A J,		
PA	(HEAR/) HEARD J.		
PA	(JIAN/) JIANG C.		
PA	(KEED/) KEDDIE J.		
PA	(ADAM/) ADAM L.		
PA	(RATC/) RATCLIFF O.		
PA	(REUB/) REUBER J L.		
PA	(RIEC/) RIECKHANN J L.		
PA	(YUGG/) YU G.		
PA	(PINE/) PINEDA O.		
XX			
PI	Pilgrim M, Creelman R, Dubell AJ, Heard J, Jiang C, Keddie J;		
PI	Adam L, Ratcliffe O, Reuber JL, Rieckmann JL, Yu G, Pineda O;		
DR	WPI: 2002-292022/73.		
DR	P-PADB: AA093178.		
PT	An isolated or recombinant polynucleotide used to produce a transgenic		
PL	plant -		
PS	Claim 4; Page 881-883; 941pp: English.		
XX			
XX	The invention relates to 1 of 232 isolated or recombinant polynucleotides		
CC	encoding an Arabidopsis thaliana transcription factor, their variants,		
CC	complements, fragments, or related polynucleotide with 31% to 95%		
CC	sequence identity, where the plant possesses an altered trait as compared		
CC	to a wild-type or reference plant, or the plant exhibits an altered		
CC	phenotype as compared to a wild-type or reference plant, or the plant		
CC	exhibits ectopic expression or altered expression of one or more genes		
CC	associated with a plant trait as compared to a wild plant. Also included		
CC	are a transgenic plant comprising the polynucleotides, a computer		

[illegible]

## RESULT 11

AAH87724

ID AAH87724 standard; cDNA: 626 BP.

XX AAH87724;

XX 25-SEP-2001 (first entry)

XX Peppermint plant oil gland expressed cDNA 80.

XX Peppermint plant oil gland cell; terpenoid essential oil; resin;

XX genetic mapping; antisense suppression; recombinant expression; ss.

XX Mentha x piperita.

XX WO200153319-A1.

XX 26-JUL-2001.

XX 19-JAN-2001; 2001WO-US02567.

XX 20-JAN-2000; 2000US-0177264.

XX (CROT/) CROTEAU R. B.

XX (LANG/) LANGE B. M.

XX (WILD/) WILDUNG M. R.

XX Croteau RB, Lange BM, Wildung MR;

XX WPI: 2001-488706/53.

XX New nucleic acid molecules corresponding to mRNA molecules expressed in

XX peppermint oil glands for enhancing expression of plant oil gland cell

XX proteins -

XX Claim 1; Page 108; 251pp; English.

XX The invention relates to nucleic acid molecules (AAH87645-AAH88116) that

XX correspond to all or part of a mRNA molecule expressed in plant oil

XX gland cells, especially peppermint and plant oil glands that produce

XX terpenoid essential oils and resins. The nucleic acids are useful for

XX genetically mapping a plant genome for genes expressed in plant oil

XX gland cells and to suppress (for example by antisense suppression) or

XX enhance their expression (for example by genetically transforming a

XX plant cell with a replicable expression vector that expresses one or more

XX proteins naturally expressed in plant oil gland cells). The nucleic acids

XX are also useful for recombinant expression of plant oil gland proteins

XX required for terpenoid essential oil and/or resin production in bacterial

XX and/or yeast cells.

XX Sequence 626 BP; 187 A; 158 C; 163 G; 118 T; 0 other;

XX Alignment Scores:

XX Pred. No.: 2.16e-45 Length: 626

XX Score: 495.00 Matches: 98

XX Percent Similarity: 69.59% Conservative: 21

XX Best Local Similarity: 57.31% Mismatches: 23

XX Query Match: 43.50% Gaps: 29

XX DB: 22 Gaps: 4

XX US-10-021-811-36 (1-206) x AAH87724 (1-626)

DB 236 AACTCTGGCTGCTCAGCAGCGCCTCAACAGAACTGGAAGAGCTGACAGCTGAGATGG 295

QY 61 LeuasnTyrLeuArgProAspValArgArgGlyAsnIleThrProGlnGlnLeu 80

DB 296 CTGAACATCTCCGCCCGCCGATGTCGACGTCGACATCTCTTCAAGAGCAGCTTGG 355

QY 81 IleMetGluLeuHisAlaValTyrGlyAsnArgTyrSerIleAlaValHisLeuPro 100

DB 356 ATTCTGACCTCATCTCTGATGGGCGAACAGSTGGTCGAAATCCGCGACATCTGCCG 415

QY 101 GlyArgThrAspAsnGlnIleValAsnTyrTyrArgTyrArgIleGlnHisIleVal 120

DB 416 GGAAGAACTGATCGAATAAAGAACTACTGAGAACAAAGATGCA 463

QY 121 GlnAlaGluAsnPhenGlnGlnInSerSerAsnAsnSerGluLeuAsnAspHisGlnAla 140

DB 464 -----AACATGCGAAGCAGCTCAAAAT-----GTGACGTGCACAGCAGCA 502

QY 141 SerThrSerHisVal-----SerThrMet 148

DB 503 TTCAGAGACCCATCCCTACCTTGGATGCCCTAGCGCTGTGAGAGAAATCCAAAGCACTTC 562

QY 148 TalagluProMetGluMetTyrSerProPro 158

DB 563 TGCACTGCTCTGCTCGCTCGCTCGCTCG 593

## RESULT 12

AAC57194

ID AAC57194 standard; DNA: 524 BP.

XX AAC57194;

XX 25-JAN-2001 (first entry)

XX Eucalyptus grandis transcription factor DNA sequence #700.

XX Plant: transcription factor; gene expression: eucalyptus; pine; acacia;

XX poplar; sweetgum; teak; mahogany; bZIP; G-box binding factor;

XX basic helix-loop-helix zipper; homeotic; homeodomain; homeobox; MADS;

XX homeodomain zipper; LIM domain; AP2; FWRB5; zinc finger domain;

XX type 2 Cys2His2; CCAT box element; MYB; ss.

XX Eucalyptus grandis.

XX WO200053724-A2.

XX 14-SEP-2000.

XX 09-MAR-2000; 2000WO-US06112.

XX 11-MAR-1999; 99US-0266513.

XX 18-AUG-1999; 99US-0149485.

XX (GENE-) GENESIS RES &amp; DEV CORP LTD.

XX (FLET-) FLETCHER CHALLENGE FORESTS LTD.

XX Wood M, McGrath A, Shenk MA, Glenn M;

XX WPI: 2000-579369/54.

XX New isolated polynucleotide encoding a plant transcription factor for

XX producing a plant e.g. a woody plant, preferably eucalyptus or pine,

XX having modified gene expression or modified activity of a polypeptide

XX Claim 1; Pages 614-615; 747pp; English.

XX The present invention relates to novel plant transcription factors from

XX Eucalyptus grandis or Pinus radiata. The present sequence is the coding

XX sequence for one such transcription factor. The transcription factor may

XX be used to produce a plant having modified gene expression such as a

XX woody plant e.g. a eucalyptus, pine, acacia, poplar, sweetgum, teak, or

XX mahogany species or to modify the activity of a polypeptide in a plant.



OY 153 j3UMeTtyrSerProfcyStyrGlnGlyMetLeuGluProPhseThrGlnPhe 171  
Db 535 AACGACCGAGTCGCG-----GTCGACCGCGGTTTCGTCGAATTC 573  
RESULT 14  
AAC44894  
ID AAC44894 standard; DNA; 959 BP.  
AC AAC44894;  
XX  
XX 18-OCT-2000 (first entry)  
DT  
XX  
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 44531.  
XX  
XX Hybridisation assay; genetic mapping; gene expression control;  
KW protein identification; signal transduction pathway;  
KW metabolic pathway; promoter; termination sequence; ss.  
XX  
XX Arabidopsis thaliana.  
XX  
XX EP1033405-A2.  
PD 06-SEP-2000.  
XX  
XX 25-FEB-2000; 2000EP-0301439.  
PF  
XX 25-FEB-1999; 99US-0121825.  
PR 05-MAR-1999; 99US-0123180.  
PR 09-MAR-1999; 99US-0123548.  
PR 23-MAR-1999; 99US-0125788.  
PR 25-MAR-1999; 99US-0126264.  
PR 29-MAR-1999; 99US-0126785.  
PR 01-APR-1999; 99US-0127462.  
PR 06-APR-1999; 99US-0128234.  
PR 08-APR-1999; 99US-0128714.  
PR 16-APR-1999; 99US-0129845.  
PR 19-APR-1999; 99US-0130077.  
PR 21-APR-1999; 99US-0130449.  
PR 23-APR-1999; 99US-0130510.  
PR 28-APR-1999; 99US-0130891.  
PR 30-APR-1999; 99US-0131449.  
PR 30-APR-1999; 99US-0132048.  
PR 04-MAY-1999; 99US-0132407.  
PR 05-MAY-1999; 99US-0132484.  
PR 06-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.  
PR 06-MAY-1999; 99US-0132487.  
PR 07-MAY-1999; 99US-0132863.  
PR 11-MAY-1999; 99US-0134286.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.

PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139753.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139839.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 02-JUL-1999; 99US-0142154.  
PR 06-JUL-1999; 99US-0142055.  
PR 08-JUL-1999; 99US-0142390.  
PR 09-JUL-1999; 99US-0142803.  
PR 12-JUL-1999; 99US-0142920.  
PR 13-JUL-1999; 99US-0142977.  
PR 14-JUL-1999; 99US-0143542.  
PR 15-JUL-1999; 99US-0143624.  
PR 16-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 19-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 21-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144881.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 06-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.





XX (GENE-) GENESIS RES. & DEV. CORP. LTD.  
 PA (FLEET-) FLETCHER CHALLENGE FORESTS LTD.  
 XX Wood M, McGrath A, Shenk MA, Glenn M;  
 PI WPI: 2000-579369/54.  
 XX

DR New isolated polynucleotide encoding a plant transcription factor for  
 PT producing a plant e.g. a woody plant, preferably eucalyptus or pine,  
 PT having modified gene expression or modified activity of a polypeptide  
 PT

XX Claim 1: Page 120; 747pp; English.  
 PS

XX The present invention relates to novel plant transcription factors from  
 CC Eucalyptus grandis or Pinus radiata. The present sequence is the coding  
 CC sequence for one such transcription factor. The transcription factor may  
 CC be used to produce a plant having modified gene expression such as a  
 CC woody plant e.g. a eucalyptus, pine, acacia, poplar, sweetgum, teak, or  
 CC mahogany species or to modify the activity of a polypeptide in a plant.  
 CC The transcription factors of the present invention are members from the  
 CC following families of regulatory proteins: bZIP, bZIP family of G-box  
 CC binding factors, basic helix-loop-helix zipper, LIM domain, AP2  
 CC homeotic/homeodomain/homeobox/MADS, homeodomain zipper, LIM domain, AP2  
 CC and EREBS, zinc finger domains of type 2 Cys2His2, CCAAT box elements  
 CC and MYB.  
 XX

SQ Sequence 389 BP; 112 A; 102 C; 97 G; 78 T; 0 other;

# Alignment Scores:

Pred. No.:	1,75e-37	Length:	389
Score:	421.00	Matches:	81
Percent Similarity:	88.17%	Conservative:	1
Best Local Similarity:	87.10%	Mismatches:	9
Query Match:	36.99%	Indels:	2
DB:	21	Gaps:	1

US-10-021-811-36 (1-206) x AAC56152 (1-389)

QY 1 MetAspLysLys-----GInGInCysLysThrSerGlnAspProGluValArgLysGly 18  
 Db 44 ATGGACAAGAAAGCCGACGACGACAGTGTAGTCCCAAGATGTGCGAGGTGAGAAAGGG 103  
 QY 19 ProTrpThrMetGluGluAspLeuIleLeuMetAsnTyrIleAlaAsnHisGlyGluGly 38  
 Db 104 CCGTGGACGATGAGAGAGATCTCATCTCATCACTACATAGGATCAAGGGAAGGC 163  
 QY 39 ValTrpAsnSerLeuAlaLysAlaLysGlyLeuLysArgAsnGlyLysSerCysArgLeu 58  
 Db 164 AGTTGGAACTCCCTAGCCCAAGCTGCTGCTAAACGTACCGGAGAGAGTGTCTGGGCTC 223  
 QY 59 ArgTrpLeuAsnTyrLeuArgProAspValArgArgGlyAsnIleThrProGluGluGln 78  
 Db 224 CGGTGGCTGAACTATCTGCGACCGGAGTCCGGAGAGCAACATCACTACTAGAGGACAG 283  
 QY 79 LeuLeuIleMetGluLeuHisAlaLysTrpClyAsnArg 91  
 Db 284 CTCTGATCATGGAATGCAATGCCAAGTGGGAAACAGG 322

Search completed: February 18, 2003, 10:37:54  
 Job time : 223 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: February 18, 2003, 10:31:30 ; Search time 1688 Seconds  
(without alignments)  
1976.463 Million cell updates/sec

Title: US-10-021-811-36  
Perfect score: 1138  
Sequence: 1 MDKKQCKTQSDPEVRKGPV.....NNINYSMEDSWSMQLINGD 206

Scoring table:  
BIOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 16154066 segs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL-frame+ .p2n.model -DEV-xlh  
-O/cgn2\_1/USPTO.spool/US10021811/runat\_12022003\_090133\_9566/app-query.fasta\_1.391  
-DB-EST -QPMT-fastlap -SUFFIX-p2n.rst -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0  
-UNITS-bits -START=1 -END=1 -MATRIX-biosum62 -TRANS-human40.cdi -LIST=45  
-DOALIGN=200 -THR SCORE=pcct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE-LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000  
-USER=US10021811\_ECGN\_1\_1\_899\_Eftunat\_12022003\_090133\_9566 -NCPD=6 -ICPD=3  
-NO\_XLPXY -NO\_MMAP -LARGEQUERY -NDC\_SCORES=0 -WAIT -LONGLOG -DEV-TIMEOUT=120  
-YGAPOP=10 -YGAPEXT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -YGAPOP=6 -YGAPEXT=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :  
EST:\*  
1: em\_estda:\*  
2: em\_estlum:\*  
3: em\_estlin:\*  
4: em\_estlu:\*  
5: em\_estov:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_hlc:\*  
9: qb\_est1:\*  
10: qb\_est2:\*  
11: qb\_hlc:\*  
12: qb\_est3:\*  
13: qb\_est4:\*  
14: qb\_est5:\*  
15: em\_estfun:\*  
16: em\_estom:\*  
17: qb\_gss:\*  
18: em\_gss\_hum:\*  
19: em\_gss\_liv:\*  
20: em\_gss\_pln:\*  
21: em\_gss\_vtl:\*  
22: em\_gss\_fun:\*  
23: em\_gss\_mus:\*  
24: em\_gss\_mus:\*  
25: em\_gss\_other:\*  
26: em\_gss\_pro:\*  
27: em\_gss\_rod:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	% Match	Length	DB ID	Description
1	991	87.1	562	14	BM732121 sal13b11..
2	859.5	75.5	635	14	BM732121 sal13b11..
3	851	74.8	501	13	BM527606 BM527606
4	846	74.3	782	10	BE658316 BE658316
5	816	71.7	552	13	BM527774 BM527774
6	813.5	71.5	560	14	BM732539 BM732539
7	808	71.0	669	12	BM457971 BM457971
8	796	69.9	682	14	BM446831 BM446831
9	795.5	69.9	568	13	BM527664 BM527664
10	784	68.9	523	9	A1930997 A1930997
11	784	68.9	688	13	BM527897 BM527897
12	779.5	68.5	544	13	BM527508 BM527508
13	774.5	68.1	542	13	BM528383 BM528383
14	724	63.6	681	9	A1486576 A1486576
15	709.5	62.3	476	10	BE057370 BE057370
16	673.5	59.2	592	10	AM928296 AM928296
17	671	59.0	558	10	BE324639 BE324639
18	670.5	58.9	470	10	AM156542 AM156542
19	657	57.7	409	12	BE804790 BE804790
20	654.5	57.5	425	10	AM423958 AM423958
21	642	56.4	503	14	BM104458 BM104458
22	639	56.2	500	9	A1897784 A1897784
23	639	56.2	519	9	A1897681 A1897681
24	637.5	55.0	502	14	BM106505 BM106505
25	627.5	55.1	447	13	BI273011 BI273011
26	621	54.6	395	12	BE805071 BE805071
27	614	54.0	450	10	BE058847 BE058847
28	601	52.8	431	10	AM459279 AM459279
29	595	52.3	586	12	BE820766 BE820766
30	593	52.1	715	10	BE658737 BE658737
31	529.5	46.5	542	9	AU238946 AU238946
32	527	46.3	397	10	AM433264 AM433264
33	526.5	46.3	727	12	BE607379 BE607379
34	522.5	45.9	676	14	BM013207 BM013207
35	514	45.2	588	12	BE649523 BE649523
36	513.5	45.1	766	10	BE659054 BE659054
37	513.5	45.1	635	10	AM032656 AM032656
38	513.5	45.1	690	14	BM147546 BM147546
39	511	44.9	605	13	BM233398 BM233398
40	510	44.8	370	12	BE642441 BE642441
41	510	44.8	371	9	A1487923 A1487923
42	508	44.6	502	10	BE459188 BE459188
43	506.5	44.5	566	13	BI424418 BI424418
44	506	44.5	510	12	BE325282 BE325282
45	501	44.0	536	13	BI701275 BI701275

## ALIGNMENTS

RESULT 1  
LOCUS BM732121  
DEFINITION sal13b11.y1 Gm-c1061 Glycine max cDNA clone SOYBEAN CLONE ID:  
Gm-c1061-4678 5' similar to TR:Q39028 Q39028 ATMYB2. ;, mRNA  
ACCESSION BM732121  
VERSION BM732121  
KEYWORDS EST.  
SOURCE soybean.  
ORGANISM Glycine max  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;  
Glycine.

REFERENCE 1 (bases 1 to 562)  
 AUTHORS Shoemaker, R., Keim, P., Vodkin, L., Erpeiding, T., Corvett, V., Khanna, A., Bolla, B., Marra, M., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wylie, T., Underwood, K., Stepien, M., Theising, B., Allen, M., Bowers, Y., Person, B., Swaller, T., Gibbons, M., Page, D., Harvey, N., Schurk, R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R., Waterson, R. and Wilson, R.  
 TITLE Public Soybean EST Project  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Shoemaker R/Public Soybean EST Project  
 Public Soybean EST Project  
 Washington University School of Medicine  
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA  
 Tel: 314 286 1800  
 Fax: 314 286 1810  
 Email: est@watson.wustl.edu  
 This clone is available through: ResGen, Invitrogen Corp. 2130 South Memorial Parkway Huntville, AL 35801 for further information call: (800) 533-4363 or contact: ccl@resgen.com web site: www.resgen.com  
 Seq primer: -40RP from Glibco  
 High quality sequence stop: 421.  
 Location/Qualifiers  
 1..562  
 /organism="Glycine max"  
 /db\_xref="taxon:3847"  
 /clone="SOYBEAN CLONE ID: Gm-cl061-4678"  
 /clone\_1lb="Gm-cl061"  
 /tissue\_type="mature flowers of field grown plants"  
 /lab\_host="DH10B"  
 /note="Vector: plasmid II SK+; Site: 1: EcoRI; Site 2: XhoI; The cDNA library was constructed from mRNA isolated from mature flowers of field grown plants for the cultivar Radden. Complementary DNA was synthesized from mRNA using a primer consisting of a poly(dT) sequence with a XhoI restriction site. EcoRI adapters were ligated to the blunt-ended cDNA fragments followed by XhoI digestion. The cDNA fragments were directionally cloned into the EcoRI-XhoI restriction site of the plasmid vector. The ligated cDNA fragments were transformed into DH10B host cells (GibcoBRL). This library was constructed in the laboratory of Dr. Randy Shoemaker."

BASE COUNT 167 a 115 c 128 g 132 t  
 ORIGIN

Alignment Scores:  
 Pred. No.: 2.7e-106 Length: 562  
 Score: 991.00 Matches: 181  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 87.08% Indels: 0  
 DB: 14 Gaps: 0

US-10-021-811-36 (1-206) x BW73121 (1-562)

QY 1 MetaSPLysLysGlnGlnCysLysThrSerGlnAspProGluValArgLysGlyProTrp 20  
 |||||||  
 DB 20 ATGATGAGAGCAACAGCTGTAAAGCTCAAGATCTCGAAGTGAAGAAAGGCGCTGG 79  
 |||||||  
 QY 21 ThrMetGluGluAspLeuIleuMetAspTryIleAlaAsnHisGlyGluGlyValTrp 40  
 |||||||  
 DB 80 ACAATGGAGAAAGACTTGAATCTTGATGACATATGTCGAATCATAGGGAAAGGTTTGG 139  
 |||||||  
 QY 41 AsnSerLeuAlaLysAlaIleuLysArgAsnGlyLysSerCysArgLysLeuArgTrp 60  
 |||||||  
 DB 140 AACTCTTGGCCAAAGCTGCTGCTCAACAGTACGAAAGAGTGGCGGCTAAGCTGG 199  
 |||||||  
 QY 61 LeuAsnTryLeuArgProAspValArgArgGlyAsnIleThrProGluGluGlnLeuLeu 80  
 |||||||  
 DB 200 CTAATTAACCTCCGTCGATGTAGAAAGAGGAAATATTACACCCGAGAAACACTTTTG 259  
 |||||||  
 QY 81 IleMetGluLeuHisAlaLysTrpGlyAsnArgTrpSerLysIleAlaLysHisLeuPro 100  
 |||||||

DB 260 ATTATGAGAGCTCCACGCAAGTGGGAACAGTGGTCCAAATTTGCCAAGCATCTACT 319  
 QY 101 G1AArgThrAspAsnGluIleLysAsnTryTTPArgThrArgIleGlnLysHisLys 120  
 |||||||  
 DB 320 GGAAGGCTGATATATAGATCAAGAACTATTGGAGACAGAGTCCAGAGCAATCAAG 379  
 |||||||  
 QY 121 G1A1AgluAsnPhgInglInserSerAsnAsnSerGluIleAsnAspHisGlnAla 140  
 |||||||  
 DB 380 CAAGCTCAGAACTTTCAGCAGCAGACTAGTAATTAATTCGAGATTAATATATCACCAGCT 439  
 |||||||  
 QY 141 SerThrSerHisValSerThrMetAlaGluProMetGluMetTyrSerProCysTyr 160  
 |||||||  
 DB 440 AGCACTAGCCATTTTCACACATGCTGAGCCCATGGAATGATTTCTCCACCTGTAT 499  
 |||||||  
 QY 161 G1nGlyMetLeuGluProPheSerThrGlnPheProThrIleAsnProAspInserSer 180  
 |||||||  
 DB 500 CAAGGAATGTATAGAGCATTTTCACTCACTAGTTCCCTACAAATTAATTCGATCAATCCACT 559  
 |||||||  
 QY 181 Cys 181  
 |||  
 DB 560 TGT 562

RESULT 2  
 LOCUS B0146360 635 bp mRNA linear EST 24-APR-2002  
 DEFINITION NF047E06FL 5', mRNA sequence.  
 ACCESSION B0146360  
 VERSION B0146360.1 GI:20283419  
 KEYWORDS EST.  
 SOURCE barrel medic.  
 ORGANISM Medicago truncatula  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifoliaceae; Medicago. 1 to 635)

REFERENCE 1 (bases 1 to 635)  
 AUTHORS Torres-Jerez, I., Scott, A.D., Harris, A.R., Gonzales, R.A., Bell, C.J., Flores, R.R., Inman, J.T., Weller, J.W. and May, G.D.  
 TITLE Expressed Sequence Tags from the Samuel Roberts Noble Foundation Medicago truncatula flower library  
 JOURNAL Unpublished (2001)  
 COMMENT Contact: May GD  
 Plant Biology Division  
 The Samuel Roberts Noble Foundation  
 2510 Sam Noble Parkway, Ardmore, OK 73402, USA  
 Tel: 580 221 7391  
 Fax: 580 221 7380  
 Email: gdmay@noble.org  
 Insert Length: 635 Std Error: 0.00  
 Plate: 047 row: E column: 06  
 Seq primer: TCACACAGGAACAGCTATGAC.  
 Location/Qualifiers  
 1..635  
 /organism="Medicago truncatula"  
 /db\_xref="taxon:3880"  
 /clone="NF047E06FL"  
 /clone\_1lb="Developing flower"  
 /tissue\_type="Developing flowers"  
 /dev\_stage="Developmentally pooled. Contains a mixture of very young, developing, fully-opened flowers and flowers in early transition into pods."  
 /note="Vector: Lambda Zap; cDNA was prepared from polyA+ enriched, pooled samples of equivalent amounts of total RNA from very young, developing, fully-opened flowers and flowers transitioning into pods. The cDNA was directionally ligated into the uni-zap XR vector (Stratagene) and packaged using the GigaPack III gold packaging extracts. Phagemids containing cDNA inserts were in vivo excised from the recombinant uni-zap XR vector using Exsist helper phage and the E. coli strain X1-Blue MRF<sup>+</sup> (Stratagene). Excised plasmids were plated using SOLR cells."

BASE COUNT 229 a 121 c 117 g 168 t  
 ALIGNMENT SCORES:  
 Pred. No.: 9,86e-91 Length: 635  
 Score: 859.50 Matches: 167  
 Percent Similarity: 85.28% Conservative: 7  
 Best Local Similarity: 81.86% Mismatches: 18  
 Query Match: 75.53% Indels: 12  
 DB: 14 Gaps: 5  
 US-10-021-811-36 (1-206) x B0146360 (1-635)

QY 1 MetaspLysLysGlnGlnCysLysThrSerGlnAspProGluValArgLysGlyProTrrp 20  
 DB ATGGATTAATAA---CCATGCAACCTCATCTCAAGATCCTGAGAGTGAAGGAGGCGCATGG 93  
 QY 21 ThrMetGluGluAspLeuIleLeuMetAsnTyrIleAlaAsnHisGlyGluGlyValTrrp 40  
 DB ACCATGGAAGAACCTGTTGATTTTATTCATTTATTTGCAAAATCATGCTGAACGCTTTGG 153  
 QY 41 AsnSerLeuAlaLysAlaIleGlyLeuLysArgAsnGlyLysSerCysArgLeuArgTrrp 60  
 DB 154 AATTCCTTGAAGCAAAAGCTCTGCTTAAACGTACAGGAAAAAGTTGCGAGCTTCGATGG 213  
 QY 61 LeuAsnTyrLeuArgProAspValArgArgGlyAsnIleThrProGluGlnGluLeu 80  
 DB 214 TTTAACTACTCTGCTCAATGTTAGAGAGGAAATTTACCTGAGGAAACACTTTTG 273  
 QY 81 IleMetGluLeuHisAlaLysTrrpGlyAsnArgTrrpSerLysIleAlaLysHisLeuPro 100  
 DB 274 ATCATGGAACCTCATCTCAAGAGTGGGAAATAGCTGCTCAAAATTCGCAAGATCTTCCA 333  
 QY 101 G1AArgThrAspAsnGluIleLysAsnTyrTrrpArgIleGluLysHisIleLys 120  
 DB 334 GGAAGACACACATAGATTAAGAAATTTTGGAGAGACATAGATACAAAGACACATTTAG 393  
 QY 121 GlnAlaGlu-----AsnPhenGlnGlnSerSerAsnAsnSerGluIle 135  
 DB 394 CAAGTGTATCATCTCACCACAAATATTTCCAAACAAATGAGT-----TTAGAGATA 444  
 QY 136 AsnAsp---HisGlnAlaSerThrSer-HisValSerThrMetAlaGluProMetGluMe 154  
 DB 445 AATGATCATCATCAATCAAGTGTAGCCCAAGTTTCCCAATTTGAGACCAATGGAATTC 504  
 QY 154 TyrSerProProCysTyrGlnGlyMetLeuGluProPheSerThrGlnPheProThrI 174  
 DB 505 TTACTCTCCACCTTCATTCACCAAGAACCTTTGGAGCATTTCCAGCTCAATTTCCAAAT 564  
 QY 174 eAsnProAspGln---SerSerCysCysThrAsnAspAsnAsnIleAsnTyrTrrpSe 193  
 DB 565 CATGATCATCATCTCCTCAAGTTGTTGACAAATGACAAATACAAATATTTAGG 624  
 QY 193 rMetGluAsp 196  
 DB 625 CATGAGAGAT 634  
 RESULT 3  
 LOCUS BM527606 501 bp mRNA linear EST 19-FEB-2002  
 DEFINITION sal63906.y1 Gm-cl061 Glycine max cDNA clone SOYBEAN CLONE ID:  
 Gm-cl061-3780 5' similar to TR:Q39028 Q39028 ATMYB2. ; mRNA  
 sequence.  
 ACCESSION BM527606  
 VERSION BM527606.1 GI:18733434  
 KEYWORDS EST.  
 SOURCE soybean.  
 ORGANISM Glycine max  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;  
 Glycine  
 REFERENCE 1 (bases 1 to 501)

AUTHORS Shoemaker R., Keim P., Vodkin L., Eppelding J., Coryell V., Khanna A., Bolla B., Marita M., Hillier L., Kucoba T., Martin J., Beck C., Wylie T., Underwood K., Sleptoe M., Theising B., Allen M., Bowers Y., Person B., Swaller T., Gibbons M., Page D., Harvey N., Schuck R., Riller E., Kohn S., Shin T., Jackson Y., Cardenas M., McCann R., Waterson R. and Wilson R.  
 TITLE Public Soybean EST Project  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Shoemaker R./Public Soybean EST Project  
 Public Soybean EST Project  
 Washington University School of Medicine  
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA  
 Tel: 314 286 1800  
 Fax: 314 286 1810  
 Email: est@wustl.edu  
 This clone is available through: Resgen, Invitrogen Corp. 2130 South Memorial Parkway Huntsville, AL 35801 For further information call: (800)-533-4363 or contact: c@resgen.com web site: www.resgen.com  
 Seq primer: -40RP from Gibco  
 High quality sequence stop: 421.  
 Location/Qualifiers  
 1..501  
 /organism="Glycine max"  
 /db\_xref="taxon:3847"  
 /clone="SOYBEAN CLONE ID: Gm-cl061-3780"  
 /clone\_1db="Gm-cl061"  
 /tissue\_type="mature flowers of field grown plants"  
 /lab\_host="DH10B"  
 /note="Vector: Bluescript II SK+, Site\_1: EcoRI, Site\_2: XhoI; The cDNA library was constructed from mRNA isolated from mature flowers of field grown plants for the cultivar Raiden. Complementary DNA was synthesized from mRNA using a primer consisting of a poly(dT) sequence with a XhoI restriction site. EcoRI adapters were ligated to the blunt-ended cDNA fragments followed by XhoI digestion. The cDNA fragments were directionally cloned into the EcoRI-XhoI restriction site of the Bluescript vector. The ligated cDNA fragments were transformed into DH10B host cells (GibcoBRL). This library was constructed in the laboratory of Dr. Randy Shoemaker."

BASE COUNT 170 a 104 c 114 g 113 t  
 ORIGIN  
 ALIGNMENT SCORES:  
 Pred. No.: 6,89e-90 Length: 501  
 Score: 851.00 Matches: 157  
 Percent Similarity: 96.34% Conservative: 1  
 Best Local Similarity: 95.73% Mismatches: 6  
 Query Match: 74.78% Indels: 0  
 DB: 13 Gaps: 0  
 US-10-021-811-36 (1-206) x BM527606 (1-501)

QY 1 MetaspLysLysGlnGlnCysLysThrSerGlnAspProGluValArgLysGlyProTrrp 20  
 DB 10 ATGGATTAATAA---CCATGCAACCTCATCTCAAGATCCTGAGAGTGAAGGAGGCGCATGG 69  
 QY 21 ThrMetGluGluAspLeuIleLeuMetAsnTyrIleAlaAsnHisGlyGluGlyValTrrp 40  
 DB 70 ACGATGGAAGAACCTGATCTGATCAACGTAATTTGCAAAATCATGCGGAGAGCTTTGG 129  
 QY 41 AsnSerLeuAlaLysAlaIleGlyLeuLysArgAsnGlyLysSerCysArgLeuArgTrrp 60  
 DB 130 AATTCCTTGAAGCAAAAGCTCTGCTTAAACGTACAGGAAAAAGTTGCGGATGAGTGG 189  
 QY 61 LeuAsnTyrLeuArgProAspValArgArgGlyAsnIleThrProGluGlnGluLeu 80  
 DB 190 CTAAACCTACCTCGCTCTGATTTAGAGAGGAAATTTACACCGAGGAAACACTTTTG 249  
 QY 81 IleMetGluLeuHisAlaLysTrrpGlyAsnArgTrrpSerLysIleAlaLysHisLeuPro 100  
 DB 250 ATCATGGAACCTTCACGCAAGTGGGAAACAGTGTCTCCAAATTTGCCAAGCATCTACT 309

OY 101 G1YArGThAspaSnglUlllelySAsnTYTTPArGThArGllleGlnlyShlilelyS 120  
 |||||||  
 Db 310 GGTAGGCAATATATGATGATCAAGAACTATGGAGACCGAGATCCAGAACCATCAAC 369  
 OY 121 G1AlaGluAsnPhelGlnGlnInSerSerAsnSercGluUlleAsnAspHisGlnAla 140  
 |||||||  
 Db 370 CAAGCTGAGAACTTTCAGCAAAATTAATTAATACCTGAGATTAATGATGATCAACCA 429  
 OY 141 SerThSerHisValSerThMetAlaGluProMetGluMetYrSerProProCysTYR 160  
 |||||||  
 Db 430 AGCACTAGCAATGTTTACCATGCTGAGAACCATGAGACCTATTCCTCCACCTTTTAT 489  
 OY 161 GlnGlyMetLeu 164  
 |||||||  
 Db 490 CAAGGATGTTA 501  
 RESULT 4 BE658316 782 bp mRNA linear EST 24-MAY-2001  
 LOCUS BE658316 GM700005B10E4 Gm-r1070 glycine max cDNA clone Gm-r1070-1759 3'  
 DEFINITION mRNA sequence.  
 ACCESSION BE658316  
 VERSION BE658316.1 GI:9984208  
 KEYWORDS EST.  
 SOURCE soybean.  
 ORGANISM Glycine max  
 Glycine max  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;  
 Glycine.  
 1 (bases 1 to 782)  
 Vodka,L., Kelm,P., Shoemaker,R., Retzel,E., Khanna,A., Coryell,V.,  
 Expanding,J., Rapp,C., Shoop,E., Pardinas,J., Liu,L. and Lewin,H.  
 A Functional Genomics Program for Soybean (NSF 9872565)  
 Unpublished (1999)  
 Other\_ESTS: A1930997 corresponding to Gm-cl015-278 (5')  
 Contact: Vodka, L.O., PI, A Functional Genomics Program for  
 Soybean (NSF 9872565)  
 Lewin, H. A., Director, Keck Center for Comparative and Functional  
 Genomics  
 University of Illinois  
 Edwin R. Madigan Building, 1201 W. Gregory, Urbana, IL 61801, USA  
 Tel: (217) 244-6147  
 Fax: (217) 333-4582  
 Email: l-vodka@uiuc.edu  
 This clone is available through: Genome Systems, Inc. 4633 World  
 Parkway Circle St. Louis, Missouri 63134. For further information  
 call: (800) 430-0030 or (314) 427-3222 FAX:(888) 919-3324 or (314)  
 427-3324 or contact: clones@genomesystems.com or info@genome  
 systems.com web site: www.genomesystems.com  
 Seq primer: 5'-TTTTTTTTTTTTTTTTT(A/C/G)-3'.  
 Location/Qualifiers  
 1..782  
 /organism="Glycine max"  
 /db\_xref="taxon:3847"  
 /clone="Gm-r1070-1759"  
 /clone\_11b="Gm-r1070"  
 /note="The library Gm-r1070 is a sequence-driven, rerecked  
 set of 9,216 clones selected from cDNA libraries from  
 various tissues and stages of development of soybean that  
 represent 2,639 sequences from immature cotyledons, 1,770  
 from immature seed coats, 3,938 from flowers, and 869  
 from young pods. The 5' ESTs of the source clones from  
 the different libraries was used to select singletons, or  
 a representative of each contig, which were rerecked to  
 form library Gm-r1070. The cDNA clones of the rerecked  
 Gm-r1070 library were then sequenced at the 3' end. The  
 contig analysis to select unique genes was performed by  
 the laboratory of Ernest Retzel, Center for Computational  
 Genomics and Bioinformatics, University of Minnesota,  
 http://www.cbc.umn.edu/ResearchProjects/Soybean/index.html  
 Rerecking was performed by Genome Systems, St. Louis,

http://www.genomesystems.com, and 3' sequencing by the  
 Keck Center for Comparative and Functional Genomics,  
 University of Illinois,  
 http://www.life.uiuc.edu/biotech/Keck.html. Note: The  
 corresponding 5' EST from each clone in the Gm-r1070  
 library is listed in the 'OTHER EST' field. The detailed  
 information on the source library for each clone can also  
 be obtained by referring to the Genome Systems clone ID of  
 the original cDNA library that is also listed under  
 'OTHER EST'.  
 BASE COUNT 215 a 144 c 143 g 253 t 27 others  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 5.25e-89 Length: 782  
 Score: 846.00 Matches: 162  
 Percent Similarity: 90.61% Conservative: 2  
 Best Local Similarity: 89.50% Mismatches: 14  
 Query Match: 74.34% Indels: 3  
 DB: 10 Gaps: 2  
 US-10-021-811-36 (1-206) x BE658316 (1-782)  
 OY 29 MetAsnTYrTlLeAlaAsnHISGlyGluGlyAlaTTPAsnSerLeuAlaAlaGly 48  
 :::::::::::::::::::: |||||  
 Db 780 ATCAACTATATGCAAAATCATCTNNNNNGGTGTGANNNNCTTGCGCAAGCTGCTMNT 721  
 OY 49 LeuLYsArgAsnGlySerCysArgLeuArgTTPLeuAsnTYrLeuArgProAspVal 68  
 |||||||  
 Db 720 CTCAAACGTACCGGAGAAAGCTTCCGCGCTAAGTGCGCTTAACCTCGCTCGTATGTT 661  
 OY 69 ArgArgGlyAsnTlLeThrProGluGluLeuLleMetGluLeuHisAlaLysTrp 88  
 |||||||  
 Db 660 AGAAGGGGAAATATNNACCAGAGAACAACTTTGATCATGTGACCTTCAGCAAGG 601  
 OY 89 G1YAsnArgTrpSerLysTlLeAlaLysHISLeuProGlyArgTrpAsnGluLleLys 108  
 |||||||  
 Db 600 GGAACAGGTGTCGCAAAATTTCCCAAGCATCTACTCGTAGAGACAGATATAGATCAAG 541  
 OY 109 AsnTYrTPArGThArGllleGlnlyShlilelyGlnAlaGluAsnPhelGlnGln 128  
 |||||||  
 Db 540 AACTATTGAGAGCCAGATCCAGAACCATCAAGCAAGCAAGCAAGCAAGCAAGCAAG 481  
 OY 129 SerSerAsnSercGluUlleAsnAspHisGlnAlaSerThSerHisValSerThMet 148  
 |||||||  
 Db 480 ATAGCAATATACCTGATGATTAATGATCAAGCATAGCATAGCATGTTCTCATG 421  
 OY 149 AlaGluProMetGluMetYrSerProProCysTYrGlnGlyMetLeuLupProPheSer 168  
 |||||||  
 Db 420 GCTGAACCATGAGACCTATTCCTCCACCTTTATCAAGGAATGTTAGAGCCATTTCT 361  
 OY 169 Thr---GlnPheProThrIleAsnProAspGlnSerSerCysGlyThrAsnAspAsn 187  
 :::::::::::::::::::: |||||  
 Db 360 TCATTCACAGTCCCCACCAATTAATCTGATCAATCCAGTTGTTGTAACCAATGACAAAC 301  
 OY 188 Asn---IleAsnTYrTrpSerMetGluAspSerTrpSerMetGln-LeuLeuAsnGlyAs 206  
 |||  
 Db 300 AACGATTAAACTATTTGAGCATGAGGATATCTGTGATATGATATGATGACGGGGA 241  
 OY 206 P 206  
 Db 240 T 240  
 RESULT 5 BM527774 552 bp mRNA linear EST 19-FEB-2002  
 LOCUS BM527774 sal65903.y1 Gm-cl061 glycine max cDNA clone SOYBEAN CLONE ID:  
 DEFINITION Gm-cl061-4134 5' similar to TR:049020 049020 MYB-LIKE DNA-BINDING  
 DOMAIN PROTEIN. ; mRNA sequence.  
 ACCESSION BM527774  
 VERSION BM527774.1 GI:18733722  
 KEYWORDS EST.  
 SOURCE soybean.







```

Db 527 ACAATATATATATACACCCCTCTTACAGAGAAATTTTGGAACCATTTTCCACCTTCATCAT 586
      ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Oy 170 GlnpHeProthrIleAsnProAspGlnSerSerCysCysThrAsnAspAsnAsn1le 189
      |||||||
Db 587 CAGTTCCTCCACAAATCTCT---GATCAATCAGAGTCTCTCT-----622
Oy 190 AsnTYrTPSerMetGluAspSerTrpSerMetGlnLeuAsn 204
      |||||||
Db 623 AACTACTGAGCATGGAGATCTCTTGTCACTTCAACTACTAAAT 667
      |||||||

RESULT 8
B0146831
LOCUS B0146831 682 bp mRNA linear EST 24-Apr-2002
DEFINITION NF028606FL1051 Developing flower Medicago truncatula cDNA clone
ACCESSION B0146831
VERSION B0146831
KEYWORDS EST.
SOURCE barrel medic.
ORGANISM Medicago truncatula
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifoliaceae;
Medicago.
1 (bases 1 to 682)
Torres-Jerez, I., Scott, A.D., Harris, A.R., Gonzales, R.A., Bell, C.J.,
Flores, H.R., Iman, J.T., Weller, J.W. and May, G.D.
Expressed Sequence Tags from the Samuel Roberts Noble Foundation
Medicago truncatula flower library
Unpublished (2001)
Contact: May GD
Plant Biology Division
The Samuel Roberts Noble Foundation
2510 Sam Noble Parkway, Ardmore, OK 73402, USA
Tel: 580 221 7391
Fax: 580 221 7380
Email: gdmay@noble.org
Insert length: 682 Std Error: 0.00
Plate: 028 Row: 6 Column: 06
Seq primer: TCACACGAGAACAGCTATGAC.
Location/Qualifiers
1.682
/organism="Medicago truncatula"
/db_xref="taxon:3880"
/clone="NF028606FL"
/tissue_type="Developing flower"
/dev_stage="Developmentally pooled. Contains a mixture of
in early transition into pods."
/note="Vector: Lambda Zap; cDNA was prepared from polyA+
enriched, pooled samples of equivalent amounts of total
RNA from very young, developing, fully-opened flowers and
flowers transitioning into pods. The cDNA was
directionally ligated into the Uni-Zap XR vector
(Stratagene) and packaged using the GigaPack III Gold
packaging extracts. Phagemids containing cDNA inserts were
in vivo excised from the recombinant Uni-Zap XR vector
using ExAssist helper phage and the E. coli strain
XL1-Blue MRF' (Stratagene). Excised plasmids were placed
using SQR cells."
BASE COUNT 230 a 142 c 128 g 180 t 2 others
ORIGIN
Alignment Scores:
Pred. No.: 3.29e-83 Length: 682
Score: 796.00 Matches: 160
Percent Similarity: 81.19% Conservative: 17
Best Local Similarity: 73.39% Mismatches: 20
Query Match: 69.95% Indels: 21
DB: 14 Gaps: 7

```

```

us-10-021-811-36 (1-206) x B0146831 (1-682)
Oy 1 MetAspLysIysGlnGlnCysLysThrSerGlnAspProGluValArgLysGlyProTrp 20
      |||||||
Db 47 ATGCACAAAAA---GAGTCGAGCTTCTTCCACAGATCTCATGTAGCAAGGACCATG 103
Oy 21 ThrMetGluGluAspLeuIleLeuMetAsnTYrIleAsnHisGlyGluValTrp 40
      |||||||
Db 104 ACATGAGAAAGAGCTTGATCTTGATCACTATATTTGCCATATATGGTAGGGTTGG 163
Oy 41 AsnSerLeuAlaLysAlaIleGlyLeuLysArgAsnGlyLysSerCysArgLeuArgTrp 60
      |||||||
Db 164 AACTCTTGCTGAATCTGCTGCTTAAACGTAACCGAAGAGGTTGACAGCTTGCTGG 223
Oy 61 LeuAsnTYrLeuArgProAspValArgArgGlyAsnIleThrProGluGluLeu 80
      |||||||
Db 224 CTAACTATCTTGACCGGATGTTAGCGAGGGAATATACCTTAGGAAACAACTCTG 283
Oy 81 IleMetGluLeuHisAlaLysTrpGlyAsnArgTrpSerLysIleAlaLysHisLeuPro 100
      |||||||
Db 284 ATCATTAAGCTTATGCTAGTAGTGGGAAACAGGTGGTCCGAAATTCGAAACATCTACCG 343
Oy 101 GlyArgThrAspAsnGluIleLysAsnTYrTrpArgTrpArgIleGlnLysHisLys 120
      |||||||
Db 344 GGAAGAACCGATATGATGATGAAGAACTACTCGAGGCAAGATCCAAAGCATATCAAG 403
Oy 121 GlnAlaGluAsnPhenIn---GlnGlnSerSerAsnAsnSerGluIleAsnAspHisGln 139
      |||||||
Db 404 CAGCTGAGACTCTCAATCTCAAGAACAGGTTCTGATATTCAAATTAATGATGATAAT 463
Oy 140 -----AlaSerThrSerHisValSerThrMetAlaGluProMetG1 153
      |||||||
Db 464 AATAATATCACTTATAGGGAAGCACAGCCAAATTTCCAACTTGCTGAACTATGGA 523
Oy 153 u-----MetTYrSerProProcysTYrGlnGlyMetLeuGluProPhSerThr----- 169
      |||||||
Db 524 CACAAATATTATATATACACCCCTCTTACCAAGAAATTTTGGAACCATTTCCACCTCATCA 583
Oy 170 -GlnpHeProthrIleAsnProAspGlnSerSerCysCysThrAsnAspAsnAsn1 189
      |||||||
Db 584 TCAGTTCCTCCACAAATCTCT---GATCAATCA-----GGTTGATC 619
Oy 189 eAsnTYrTPSerMetGluAspSerTrpSerMetGlnLeuAsnGlyAsp 206
      |||||||
Db 620 TAACACTGAGCATGAGATCTCTTGCACCTCAACTCACTAATTAATGGAGAT 671
      |||||||

RESULT 9
BMS27664 568 bp mRNA linear EST 19-FEB-2002
LOCUS BMS27664
DEFINITION sal64d10.y1 Gm-c1061 Glycine max cDNA clone SOYBEAN CLONE ID:
Gm-c1061-4003 5' similar to TR:Q39028 Q39028 ATMBY2. ;, mRNA
sequence.
ACCESSION BMS27664
VERSION BMS27664
KEYWORDS EST.
SOURCE soybean.
ORGANISM Glycine max
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
Glycine.
1 (bases 1 to 568)
Shoemaker, R., Keim, P., Vodkin, L., Erpelting, J., Corvelli, V., Khanna
A., Bolla, B., Mair, M., Hiller, L., Kucaba, T., Martin, J., Beck, C.,
Wylie, T., Underwood, K., Steptoe, M., Theisling, B., Allen, M., Bowers
Y., Person, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schurk
R., Ritter, E., Kohn, S., Shih, T., Jackson, Y., Cardenas, M., McCann
R., Waterston, R. and Wilson, R.
Public Soybean EST Project
Unpublished (1999)
Contact: Shoemaker R/Public Soybean EST Project
Public Soybean EST Project
Washington University School of Medicine

```

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA  
 Tel: 314 286 1800  
 Fax: 314 286 1810  
 Email: estewatson.wustl.edu  
 This clone is available through: Resgen, Invitrogen Corp. 2130  
 South Memorial Parkway Huntville, AL 35801 For further information  
 call: (800)-533-4363 or contact: ccuteresgen.com web site:  
 www.resgen.com  
 Seq primer: -40RP from gibco  
 High quality sequence stop: 422.

## FEATURES

source

Location/Qualifiers

1..568

/organism="Glycine max"

/db\_xref="taxon:3847"

/clone="SOYBEAN CLONE ID: Gm-cl015-4003"

/clone\_lib="Gm-cl015"

/tissue\_type="Mature flowers of field grown plants"

/lab\_host="DH10B"

/note="Vector: Bluescript II SK+; Site\_1: EcoRI; Site\_2:  
 XhoI; The cDNA library was constructed from mRNA isolated  
 from mature flowers of field grown plants for the cultivar  
 Raiden. Complementary DNA was synthesized from mRNA using  
 a primer consisting of a poly(dT) sequence with a XhoI  
 restriction site. EcoRI adapters were ligated to the  
 blunt-ended cDNA fragments followed by XhoI digestion. The  
 cDNA fragments were directionally cloned into the  
 EcoRI-XhoI restriction site of the Bluescript vector. The  
 ligated cDNA fragments were transformed into DH10B host  
 cells (GibcoBRL). This library was constructed in the  
 laboratory of Dr. Randy Shoemaker."

BASE COUNT 174 a 134 c 124 g 136 t  
 ORIGIN

## Alignment Scores:

Score: 2.85e-83 Length: 568  
 Percent Similarity: 795.50 Matches: 153  
 Best Local Similarity: 87.91% Conservative: 7  
 Query Match: 69.90% Mismatches: 19  
 Indels: 3  
 Gaps: 3

US-10-021-811-36 (1-206) x BM527664 (1-568)

QY 1 Metaspyslysginglncyslysthsrclnasprrgguvalarglyglyprotrp 20  
 |||||||  
 DB 29 ATGACGACGAAGCTTGGC--AACACGCTCATATCTCGATCGAGACAAAGGGCCATGG 85  
 |||||||  
 QY 21 Thmetglugluasplesuleleumetasnryrlllealasnhsiglyuglyvaltrp 40  
 |||||||  
 DB 86 ACAATGGAAAGACTTAATCTGATCACCCTATTCACATCGGGGAAGGGCTTGG 145  
 |||||||  
 QY 41 Asnserleualalyalaaaglyleulysrrgansglysserasyarqleuargtrp 60  
 |||||||  
 DB 146 AACCTTTGGCCAAAGCGCTGGAGCTTAACGTCACCGGAAGACTTCCCGCTCCGCTCG 205  
 |||||||  
 QY 61 Leuasnryrleuargtrpaaavalaargaglyasnllrthrrprogluglucleuleu 80  
 |||||||  
 DB 206 CTAACCTACCTCCCTCCGATGTAGAGAGGGAATATTACCCCGAGGAACACCTTTG 265  
 |||||||  
 QY 81 llemetgluleuhsalalystrpcllyasnargtrpserlyllealalyhsileupro 100  
 |||||||  
 DB 266 ATCATGGAACCTTCATCAAAAGTGGGAACAGGTGCTCAAAATTTCCCAACCATCTACCC 325  
 |||||||  
 QY 101 glyargthrpsasnnglllelyshsnryrtrpargthrargllleglnlyshsillelys 120  
 |||||||  
 DB 326 GGAAGAGCTGATTAATGACATTAGAACTACGAGAGAACAGATCCAAAGCACCCTCAAG 385  
 |||||||  
 QY 121 Glnala---Gluasnphglnclnglnseraserasnsergiulleasnspshisgln 139  
 |||||||  
 DB 386 CAAGCTTCAGAGAGCTTCAGCAAGAGAGTAGTAATTCTGAGATAATTATCATCCCA 445  
 |||||||  
 QY 140 Alaserthrserhsvalserthmetlacluproketglumetgtyrserprotrp 159  
 |||||||

DB 446 GCTTCAGCTAGCAAGCTGTCACCATGGCGAGCCCATAGAAACCTATTCTCCACCCAGT 505  
 QY 160 Tyrnglmglymetleugluapropheserthrglnpneprohrlleasnproaspinsler 179  
 |||||||  
 DB 506 TATCAAGGAATGTAGTATTCATTCATTCAGTCCCAACA---AATCCATCATCATCT 562  
 QY 180 SerCys 181  
 |||||||  
 DB 563 ACTGT 568

## RESULT 10

A1930997

LOCUS

DEFINITION

Gm-cl015-278 5' similar to TR:Q39028 Q39028 ATMYB2. ;, mRNA

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

JOURNAL

COMMENT

JOURNAL

COMMENT

JOURNAL

COMMENT

JOURNAL

COMMENT

JOURNAL

COMMENT

JOURNAL

COMMENT

JOURNAL

COMMENT

JOURNAL

COMMENT

JOURNAL

COMMENT

JOURNAL

COMMENT

JOURNAL

COMMENT

JOURNAL

COMMENT

JOURNAL

COMMENT

JOURNAL

COMMENT

JOURNAL

COMMENT

JOURNAL

COMMENT

JOURNAL

COMMENT

JOURNAL

COMMENT

JOURNAL

COMMENT

JOURNAL

COMMENT

JOURNAL

COMMENT

JOURNAL

COMMENT

JOURNAL

COMMENT

JOURNAL

COMMENT

JOURNAL

COMMENT

JOURNAL

COMMENT

JOURNAL

COMMENT

JOURNAL

Alignment Scores: 5.68e-82 Length: 523

BASE COUNT 172 a 124 c 111 g 115 t 1 others

## ORIGIN

Expendig."

constructed by Dr. Randy Shoemaker and Dr. John

pluuescript vector. The ligated cDNA fragments were

transformed into XhoI-Gold host cells. This library was

cloned into the EcoRI-XhoI restriction site of the

blunt-ended cDNA fragments were directionally

ligated to the blunt-ended cDNA fragments followed by

(dT) sequence with a XhoI restriction site. EcoRI adapters

synthesized from mRNA using a primer consisting of a poly

XR cDNA library construction kit. Complementary DNA was

library was prepared using the Stratagene pluuescript II

from mature flowers of field grown plants. The cDNA

XhoI. This cDNA library was constructed from mRNA isolated

/lab\_host="XhoI-Gold"

/tissue\_type="Mature flowers, field grown plants"

/clone\_lib="Gm-cl015"

/clone="GENOME SYSTEMS CLONE ID: Gm-cl015-278"

/db\_xref="taxon:3847"

/organism="Glycine max"

Location/Qualifiers

1..523

High quality sequence stop: 422.

Seq primer: -40RP from gibco

call: (800)-533-4363 or contact via email: ccuteresgen.com

South Memorial Parkway Huntville, AL 35801 For further information

This clone is available through: Resgen, Invitrogen Corp. 2130

Email: estewatson.wustl.edu

Tel: 314 286 1800

Fax: 314 286 1810

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA

Washington University School of Medicine

Public Soybean EST Project

Contact: Shoemaker R/Public Soybean EST Project

Unpublished (1999)

R., Waterston, R. and Wilson, R.

Public Soybean EST Project

EST.

soybean.

Glycine max

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;

Glycine.

1 (bases 1 to 523)

Shoemaker, R., Kelm, P., Vodkin, L., Erpelting, J., Corvett, V., Khana

, A., Bolla, B., Matra, M., Hillier, L., Kucaba, I., Martin, J., Beck, C.,

Wyllie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers

, Y., Person, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schurk

, R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann

, R., Waterston, R. and Wilson, R.

Score: 784.00 Matches: 146  
 Percent Similarity: 94.848  
 Best Local Similarity: 94.198  
 Query Match: 68.898  
 DB: 9  
 Gaps: 0

US-10-021-811-36 (1-206) x A1930997 (1-523)

QY 1 MetasplyslsglncglnslystrSerGlnaspProgluValarglyslProtrp 20  
 DB 58 ATGGATTAACCACTGTCACACGCTCAAGATCCCGAAGTGAAGAACGACCTGG 117  
 QY 21 ThmetcgluaspLeuileuMetasnTyrlleAlasnHsglygluValargly 40  
 DB 118 ACGATGAAATAGACTTGTGATCTGATCACTATATTGCAAAATCATGGGAGGTGG 177  
 QY 41 AsnSerleuAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAla 60  
 DB 178 AATTCCTTGGCCAAAGCTGCTGCTCAAAACCTACCGAAAGAGTGGCGCTAAGGTGG 237  
 QY 61 LeuasnTyrlleuargProaspValargArgGlyAsnHleThProgluGlulnleu 80  
 DB 238 CTAACTACTCCGCTCTGATGTAGAAAGGGAATATTACACCGAGAACACTTTTG 297  
 QY 81 IlewtcgluLeuHsAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAla 100  
 DB 298 ATCATGGAGCTTCACGCAAAAGTGGGAAACAGTGGTCCAAATGGCCAAACATCTACT 357  
 QY 101 GlyArgThrAspAsnGluileuLeuMetasnTyrlleAlasnHsglygluValargly 120  
 DB 358 GGTGGGACATATATGATGATCAAGACATATGAGAGCAGACGATCCAGAGACATCAAG 417  
 QY 121 GlnAlaGluAsnPhgInglngInserSerAsnAsnSerGluileuAsnHsglnAla 140  
 DB 418 CAACCTGAGAACTGTCAGCAACATCAGCAATACTGAGATTAATGATCACCAGCT 477  
 QY 141 SerThrSerHisValSerThrMetAlaGluProMetGluMetLys 155  
 DB 478 AGCACTAGCCATGTTCTTACCATGCTGAACCCATGAGACCTAT 522

RESULT 11 688 bp mRNA linear EST 18-JUL-2001  
 LOCUS B1272897  
 DEFINITION NF091A12FL1088 Developing flower Medicago truncatula cDNA clone  
 ACCESSION NF091A12FL 5', mRNA sequence.  
 VERSION B1272897  
 KEYWORDS EST.  
 SOURCE barrel medic.  
 ORGANISM Medicago truncatula  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifoliales;  
 Medicago.  
 1 (bases 1 to 688)  
 Torres-Jerez, I., Scott, A.D., Harris, A.R., Gonzales, R.A., Bell, C.J.,  
 Flores, H.R., Iman, J.T., Weller, J.W. and May, G.D.  
 Expressed Sequence Tags from the Samuel Roberts Noble Foundation  
 Medicago truncatula flower library  
 Unpublished (2001)  
 Contact: May GD  
 Plant Biology Division  
 The Samuel Roberts Noble Foundation  
 2510 Sam Noble Parkway, Ardmore, OK 73402, USA  
 Tel: 580 221 7391  
 Fax: 580 221 7380  
 Email: gmay@noble.org  
 Insert Length: 688 Std Error: 0.00  
 Plate: 091 row: A column: 12  
 Seq primer: TCACACAGGAAACAGCTATGAC.  
 Location/Qualifiers  
 1..688  
 /organism="Medicago truncatula"

REFERENCE  
 AUTHORS  
 TITLE  
 JOURNAL  
 COMMENT

FEATURES  
 source

/db\_xref="taxon:3880"  
 /clone="NF091A12FL"  
 /clone\_lib="Developing flower"  
 /tissue\_type="Developing flowers"  
 /dev\_stage="Developmentally pooled. Contains a mixture of  
 very young, developing, fully-opened flowers and flowers  
 in early transition into pods."  
 /note="Vector: lambda Zap. cDNA was prepared from polyA+  
 enriched, pooled samples of equivalent amounts of total  
 RNA from very young, developing, fully-opened flowers and  
 flowers transitioning into pods. The cDNA was  
 directionally ligated into the Uni-Zap XR vector  
 (Stratagene) and packaged using the GigaPack III Gold  
 packaging extracts. Phagemids containing Uni-Zap XR inserts were  
 in vivo excised from the recombinant Uni-Zap XR vector  
 using Exsist helper phage and the E. coli strain  
 XL1-Blue MRF' (Stratagene). Excised plasmids were plated  
 using SOLR cells."

BASE COUNT 235 a 144 c 125 g 182 t 2 others  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 8,61e-82 Length: 688  
 Score: 784.00 Matches: 158  
 Percent Similarity: 81.578 Conservative: 19  
 Best Local Similarity: 72.818 Mismatches: 20  
 Query Match: 68.898 Indels: 21  
 DB: 13 Gaps: 7

US-10-021-811-36 (1-206) x B1272897 (1-688)

QY 1 MetasplyslsglncglnslystrSerGlnaspProgluValarglyslProtrp 20  
 DB 49 ATGGACAAAAA---GAGTGCAGTCTCTCAAGATCCCGAAGTGAAGAACGACCATGG 105  
 QY 21 ThmetcgluaspLeuileuMetasnTyrlleAlasnHsglygluValargly 40  
 DB 106 ACAATGGAAGACACTGATCTGATCAACTATATTGCCAAATCATGGGAGGTGG 165  
 QY 41 AsnSerleuAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAla 60  
 DB 166 AACTCCTTGGCTAAATGCTGCTGCTTAAACGACGGAAGAGTGGCGCTGGGG 225  
 QY 61 LeuasnTyrlleuargProaspValargArgGlyAsnHleThProgluGlulnleu 80  
 DB 226 CTAACTATCTTCGACCGGATGTTAGACGAGGAAATATTACCTGAGAACACTCTTG 285  
 QY 81 IlewtcgluLeuHsAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAla 100  
 DB 286 ATCATTTGACCTCATGCTAGTGGGAAACAGTGGTCCAAATGGCCAAACATCTACCG 345  
 QY 101 GlyArgThrAspAsnGluileuLeuMetasnTyrlleAlasnHsglygluValargly 120  
 DB 346 GGAAGAACCCATATATGATTAAGAACTACTGAGAGCAAGGATCCAAAGCATATCAAG 405  
 QY 121 GlnAlaGluAsnPhgInglngInserSerAsnAsnSerGluileuAsnHsglnAla 139  
 DB 406 CAACCTGAGAACTGCAATCTCAAGAACAGTTGCGATATTCAATTAATGATGATAT 465  
 QY 140 -----AlaserThrSerHisValSerThrMetAlaGluProMetGlu 153  
 DB 466 AATAATATCAACTTATAGAACACCAAGCAATTCACATCTCAACCTATGAGC 525  
 QY 154 -----MetTySerProProCysTyrgInglMetLeuGluProPheserThr 169  
 DB 526 ACAATATATATATACCACTCTTACCAAGAAATTTGGAAACATTTCCACCTCATCAT 585  
 QY 170 GlnPheProThrIleasnProaspGlnSerSerCysCysThrAsnAspAsnHs 189  
 DB 586 CAGTTCCTCAAACTCT---GATCATCATGAGTGTCT----- 621  
 QY 190 AsnTyTrpSerMetGluAspSerTrpSerMetGlnleuAsnGlyasp 206





FEATURES	source
116	3hlygHis1IleYsglnAlaIguaspheGlnGlnGlnSerAsnAsnSergLutIle 135
Oy	116 3hlygHis1IleYsglnAlaIguaspheGlnGlnGlnSerAsnAsnSergLutIle 135
Db	309 CAGAGACACATTTAGACAGACAGACACATCATGACACAGAGAGCAAAATATT 368
Oy	136 AsnAspHisGlnAlaSerThrSerHisValSerThrMet-----AlaGluPrometGlu 153
Db	369 GATCATCAACAGAGAAAGAACTACTAGTCAATATCATCACTGGTCAAGCAGCAATACATGAG 428
Oy	154 MetTyrSerProProCysTyrGlnGlyMetLeuGluProPheSerThrGlnPheProThr 173
Db	429 ACTTACTCTCCAACTTATCATCAATGAAATTTGGAC-----ACTAATTTT 473
Oy	174 IleAsnProAspGlnSerSerCysCysThrAsnAspAsnAsnAlleAsnTyrTrpSer 193
Db	474 -----CAAGCTCAAAATTTTCTCAATGAAACAAATGAC-----ACATGTGGAGC 518
Oy	194 MetGlnAspSerTrpSerMetGlnLeuLeuAsnGlyAsp 206
Db	519 ATGGAGGATATCTGCTCATGCAATGCTTAATGAGCAT 557
RESULT 15	
BE057370	
LOCUS	476 bp mRNA linear EST 03-DEC-2001
DEFINITION	sm02c054.y1 Gm-cl015 glycine max cDNA clone GENOME SYSTEMS CLONE ID: Gm-cl015-8363 5' similar to TR:049020 049020 MYB-LIKE DNA-BINDING DOMAIN PROTEIN. ; mRNA sequence.
ACCESSION	BE057370
VERSION	BE057370.1 GI:8401736
KEYWORDS	EST.
SOURCE	soybean.
ORGANISM	Glycine max Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eustosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.
REFERENCE	1 (bases 1 to 476)
AUTHORS	Shoemaker, R., Keim, P., Vodkin, L., Erpelting, J., Coryell, V., Khana, A., Bolla, R., Marra, M., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wylie, T., Underwood, K., Stepien, M., Theising, B., Allen, M., Bowers, T., Person, B., Swaller, T., Gibbons, M., Page, D., Harvey, N., Schurk, R., Ritters, E., Kohn, S., Shin, T., Jackson, I., Cardenas, M., McCann, R., Waterston, R. and Wilson, R.
TITLE	Public Soybean EST Project
JOURNAL	Unpublished (1999)
COMMENT	Contact: Shoemaker R/Public Soybean EST Project Public Soybean EST Project Washington University School of Medicine 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA Tel: 314 286 1800 Fax: 314 286 1810 Email: est@wustl.edu This clone is available through: Resgen, Invitrogen Corp. 2130 South Memorial Parkway Huntsville, AL 35801 For further information call: (800)-533-4363 or contact via email: ccu@resgen.com Insert Length: 909 Std Error: 0.00 Seq primer: -40RP from Gibco High quality sequence stop: 420. Location/Qualifiers 1..476

```

1. .476
/organism="Glycine max"
/db_xref="taxon:3847"
/clone="GENOME SYSTEMS CLONE ID: Gm-cl015-8383"
/clone_id="Gm-cl015"
/tissue_type="Mature flowers, field grown plants"
/lab_host="X110-Gold"
/note="Vector: pBluescript II XR; Site.1: EcoRI; Site.2:
XhoI; This cDNA library was constructed from mRNA isolated
from mature flowers of field grown plants. The cDNA
library was prepared using the Stratagene pBluescript II
XR cDNA library construction kit. Complementary DNA was
synthesized from mRNA using a primer consisting of a poly
(dT) sequence with a XhoI restriction site. EcoRI adapters
were ligated to the blunt-ended cDNA fragments followed by

```

BASE COUNT	145 a	113 c	116 g	102 t	
ORIGIN	Erpeliding.*				XhoI digestion. The cDNA fragments were directionally cloned into the EcoRI-XhoI restriction site of the pBluescript vector. The ligated cDNA fragments were transformed into XL10-gold host cells. This library was constructed by Dr. Randy Shoemaker and Dr. John
Alignment Scores:	Pred. No.:	2.9e-73	Length:	476	
Score:		709.50	Matches:	133	
Percent Similarity:		90.95%	Conservative:	6	
Best Local Similarity:		86.95%	Mismatches:	13	
Query Match:		62.35%	Gaps:	1	
DB:		10	Indices:	1	
			Gaps:	1	
US-10-021-811-36 (1-206) x BE057370 (1-476)					
QY	9 ThSerglAspProGluValArgLysGlyProTrpHrmMetGluGluAspLeuLeu				28
Db	16 ACGCTCAGATGATCTGAAAGTGAAGAAAGGGCCATGGAACATGGAAAGAACTTATTTG				75
QY	29 MetAsnTyrIleAlaAsnHisGlyGluGlyValTyrPasnSerLeuAlaAlaGly				48
Db	76 ATCACTATATTCACATCACGGGGAAGGGCTTTGAACTTTTGCCCAAGCTGTGGGA				135
QY	49 LeuLysArgAsnGlyLysSerCysArgLeuAqTrrPLeuAsnTyrLeuArgProAspVal				68
Db	136 CTTAAAGCTACCGGAAGAAAGCTTCCGCGTCGGTGGCTAAACTGAACTCCGTCGATGTT				195
QY	69 ArgArgGlyAsnIleThrProGluGluGluLeuIleMetGluLeuHisAlaLysTyr				88
Db	196 AGAAGAGGGAATTTACACCCGAGAAACGCTTTTGATCATGAGAACTTCATGCAAAATGG				255
QY	89 GlyAsnArgTyrPserLysIleAlaLysHisLeuProGlyArgTyrAspAsnGluLys				108
Db	256 GGAATACAGTGTGTCAAAATATGCCCAAGCATTCACCCGAGAGACATGATAAGATTAAAG				315
QY	109 AsnTyrTrpArgThrArgIleGlnLysHisIleLysGlnAla---GluAsnPhgGln				127
Db	316 AACATACGGAGGACAAAGATCCAGAAACACTGCAACAAAGCTTCACAGCTCCAGCA				375
QY	128 GlnSerSerAsnAsnSerGluIleAsnAspHisGlnAlaSerThrSerHisValSerThr				147
Db	376 CAGAGTAGTATTTCTGAGATAATTTTCATCCCAAGAGCTGCACATAGCAAGTGTCCACC				435
QY	148 MetAlaGluProMetGluMetTyrSerProProGlyTyr 160				
Db	436 ATGGGCGAGCCCATAGAAACCTATATTTCCACCACTGAT 474				

```
Search completed: February 18, 2003, 11:43:25
Job time : 1691 secs
```